

Northumbria Research Link

Citation: Zhang, Meng (2007) Proteomic analysis of streptococcus pyogenes. Doctoral thesis, Northumbria University.

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/842/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>



**Northumbria
University**
NEWCASTLE



UniversityLibrary

PROTEOMIC ANALYSIS OF
STREPTOCOCCUS PYOGENES

M. ZHANG

Ph.D

2007

PROTEOMIC ANALYSIS OF
STREPTOCOCCUS PYOGENES

MENG ZHANG

A thesis submitted in partial fulfillment of
the requirements of the University of
Northumbria at Newcastle for the degree of
Doctor of Philosophy

Research undertaken in the School of
Applied Sciences

March 2007

Abstract

Streptococcus pyogenes (group A streptococcus, GAS) is a major human Gram-positive pathogen that causes infections that normally occur in the respiratory tract, the skin, the wound, the lung, the bloodstream and/or muscle tissues and result in millions of deaths every year. To cause such infections, *S. pyogenes* produces a wide range of virulence factors. The destruction of connective tissue and the hyaluronic acid therein plays an important role in pathogenesis. *S. pyogenes* was propagated in hyaluronic acid rich growth media in an attempt to create a simple biological system that could reflect some elements of the pathogenesis.

The growth of bacteria was analyzed in the hyaluronic acid rich media and control media and a proteomic approach was applied to identify those proteins that were differentially expressed by the streptococcal pathogens growing in the different media. The techniques of two dimensional gel electrophoresis and static nanospray mass spectrometry were optimized and proteome maps for *S. pyogenes* grown in both media were constructed. The differentially expressed proteins by *S. pyogenes* were identified and analyzed using bioinformatics.

Our results showed that several recognized virulence factors of *S. pyogenes* were upregulated in hyaluronic acid rich media, including the M1 protein, a collagen-like surface protein and the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase, which has been shown to play important roles in streptococcal pathogenesis. Interestingly, two hypothetical proteins of unknown function were also up-regulated and detailed bioinformatics analysis showed that at least one of these hypothetical proteins is likely to be involved in GAS pathogenesis. It was therefore concluded that this simple biological system provided a valuable tool for the identification of potential streptococcal pathogens virulence factors.

Contents

ACKNOWLEDGEMENTS.....	13
DECLARATION	14
ABBREVIATIONS	15
1 INTRODUCTION.....	18
1.1 Connective tissues.....	18
1.1.1 General connective tissues	18
1.1.2 Special connective tissues.....	20
1.1.3 Amorphous ground substance.....	21
1.2 Streptococci.....	23
1.2.1 Classification of streptococci.....	23
1.2.1.1 Haemolysis	23
1.2.1.2 Lancefield grouping	23
1.2.2 GAS and GAS disease	25
1.2.3 GAS virulence factors	27
1.2.3.1 M proteins and M-like proteins	27
1.2.3.2 The HA capsule.....	32
1.2.3.3 Lipoteichoic acid.....	35
1.2.3.4 Pyrogenic exotoxins.....	36
1.2.3.5 Fibronectin-binding proteins	37
1.2.3.6 C5a peptidase.....	38
1.2.3.7 Hyaluronate lyase.....	39
1.2.3.8 Streptokinase	40
1.2.3.9 Streptolysins.....	41
1.2.3.10 DNAase.....	43
1.3 Studies of virulence factors expression in GAS.....	44
1.3.1 Development of Genomics	44
1.3.3 Transcriptomics.....	49
1.3.4 Proteomics.....	50
1.3.5 Challenge in bioinformatics.....	52
1.4 Research Objectives	53
2 MATERIAL AND METHODS.....	55
2.1 Media.....	55
2.1.1 Luria- Bertani (LB) Broth	55
2.1.2 LB agar.....	56
2.1.3 Columbia Agar with 5 % (v/v) horse blood	56
2.1.4 Defined medium.....	56
2.1.4.1 Defined medium with glucose.....	58
2.1.4.2 Defined medium with sodium hyaluronate.....	60
2.1.5 Brain Heart Infusion Broth (BHI) and BHI with 0.5 % (w/v) sodium hyaluronate Broth (BHI+HA)	60
2.1.6 Cryogenic storage of bacterial stocks	62
2.2 Growth of bacterial.....	62
2.2.1 Bacterial Strains	62
2.2.2 Storage of bacterial stocks	62

2.2.2.1 Storage of <i>E.coli</i> TOP10	62
2.2.2.2 Storage of streptococci.....	62
2.2.3 Growth of the bacteria on the agar plates.....	62
2.2.3.1 Plating <i>E. coli</i> TOP10.....	63
2.2.3.2 Streaking GAS.....	63
2.2.4 Growth of <i>E. coli</i> TOP10 for protein extraction	63
2.2.5 Growth of <i>S. pyogenes</i> SF370 in defined media.....	64
2.2.6 Growth of streptococci in BHI and BHI + HA	64
2.3 Investigation of the expression of hyaluronate lyases by <i>S. pyogenes</i> SF370	64
2.3.1 Preparation of 'enzyme' for the assays	65
2.3.2 DNSA reducing sugar assay	65
2.3.2.1 DNSA reagents.....	65
2.3.2.2 Reactions for DNSA assays	67
2.3.3 232 nm assay	69
2.4 Investigation into the consumption of carbohydrates in <i>S. pyogenes</i> SF370 cultures	71
2.4.1 Preparation of samples for assays.....	71
2.4.2 DNSA reducing sugar assay	71
2.4.3 Analysis of the uronic acid consumption in the cultures	71
2.5 2DE of <i>E. coli</i> TOP10.....	73
2.5.1 Solutions for protein extraction from Gram-negative bacteria	73
2.5.2 Protocol for protein extraction from <i>E. coli</i> TOP10	74
2.5.3 Separation of proteins using SDS-PAGE electrophoresis	75
2.5.3.1 SDS-PAGE reagents.....	75
2.5.3.2 SDS-PAGE	77
2.5.3.3 Visualisation of protein bands and photography of SDS-PAGE gels	78
2.5.4 Separation of proteins by 2DE	79
2.5.4.1 Rehydration of Immobiline DryStrip gels	79
2.5.4.1.1 Rehydration solutions.....	79
2.5.4.1.2 Rehydration of Immobiline DryStrip pH 3-10, 18 cm with rehydration sample loading	79
2.5.4.1.3 Rehydration of Immobiline DryStrip pH 3-10, 18 cm with cup sample loading.....	80
2.5.4.2 First dimension - Isoelectric focusing (IEF).....	80
2.5.4.2.1 Preparation of the Immobiline DryStrip Kit	80
2.5.4.2.2 Preparation for IEF with rehydration loaded Immobiline DryStrips	81
2.5.4.2.3 Preparation for IEF for cup loaded Immobiline DryStrips	81
2.5.4.2.4 Running IEF for Immobiline DryStrip pH 3-10, 18cm.....	81
2.5.4.3 Equilibration of Immobiline Drystrip gels	83
2.5.4.3.1 Equilibration solutions	83
2.5.4.3.2 Equilibration.....	85
2.5.4.4 Second dimension - SDS-PAGE	86
2.5.4.4.1 SDS-PAGE reagents	86
2.5.4.4.2 Casting of SDS-PAGE gel.....	87
2.5.4.4.2.1 Assembling the glass plate sandwiches.....	87
2.5.4.4.2.2 Casting the gels	87

2.5.4.4.3 Set up for SDS-PAGE	88
2.5.4.4.4 Running SDS-PAGE	89
2.5.4.5 Visualisation and photography of 2DE gels.....	89
2.6 2DE of streptococci.....	89
2.6.1 Preparation of proteins from streptococci for 2DE	89
2.6.1.1 Solutions for protein extraction	89
2.6.1.2 Isolation of cell-associated proteins from <i>S. pyogenes</i> SF370	90
2.6.1.3 Isolation of secreted proteins from <i>S. pyogenes</i> SF370 culture supernatant.....	90
2.6.1.4 Treatment with a 2-D Clean-Up kit.....	91
2.6.1.5 Determination of protein concentration.....	91
2.6.1.5.1 Bradford method (Bradford, 1976)	92
2.6.1.5.1.1 Bradford assay reagents	92
2.6.1.5.1.2 Bradford assay.....	92
2.6.1.5.2 2-D Quant Kit (Appendix A4)	93
2.6.1.5.2.1 Reagents for 2-D Quant Kit (Amersham Biosciences).	93
2.6.1.5.2.2 Assay using 2-D Quant Kit	93
2.6.2 Separation of proteins from streptococci using 2DE	95
2.6.2.1 Rehydration of Immobiline Drystrip gels	95
2.6.2.2 IEF for separation of proteins from streptococci.....	95
2.6.3 Equilibration of Immobiline DryStrip gels.....	95
2.6.4 SDS-PAGE for separation of proteins from streptococci.....	97
2.6.4.1 SDS-PAGE reagents.....	97
2.6.4.2 Set up and running of SDS -PAGE.....	97
2.6.5 Visualisation of 2DE gels	97
2.6.5.1 Colloidal Coomassie Blue staining reagents	97
2.6.5.2 Colloidal Coomassie Blue staining and destaining.....	98
2.7 Analysis of 2DE gels using PDQuestTM v 6.2.1 software.....	98
2.7.1 Image acquisition.....	100
2.7.2 Image sizing and filtering	100
2.7.2.1 Images sizing.....	100
2.7.2.2 Images filtering.....	100
2.7.3 Detecting and editing spots.....	101
2.7.3.1 Detecting spots	101
2.7.3.2 Editing spots	101
2.7.3.2.1 Adding and removing spots.....	101
2.7.3.2.2 Combining spots and changing spot boundaries	102
2.7.4 Spots comparison and matching	102
2.7.4.1 Creating a matchset	102
2.7.4.2 Matching spots.....	103
2.7.5 Data analysis	103
2.8 Protein digestion for MS analysis	104
2.8.1 Digestion of protein in solution	104
2.8.1.1 Reagents for digestion	104
2.8.1.2 Tryptic digestion.....	105
2.8.2 In-gel protein digestion	105
2.8.2.1 Reagents for in-gel protein digestion.....	105
2.8.2.2 In-gel digestion for colloidal Coomassie blue staining.....	106

2.8.3 Desalting using C ₁₈ ZipTip® (Millipore).....	107
2.8.3.1 Solutions for desalting.....	107
2.8.3.2 Desalting.....	107
2.9 MS analysis with LCQ Advantage (ThermoElectron)	108
2.9.1 Static nanospray MS set up.....	108
2.9.1.1 Static nanospray source.....	108
2.9.1.2 Installation of the static emitter	108
2.9.1.3 Loading sample into a static nanospray emitter	108
2.9.2 Tuning with the static NSI source.....	110
2.9.2.1 Tuning solutions	110
2.9.2.2 Tuning	110
2.9.3 Analysis of salts and detergent tolerance of the static NSI source	111
2.9.3.1 Preparation of protein digests for analysis	111
2.9.3.2 Collecting data	112
2.9.4 MS analysis for protein identification using the static nanospray	112
2.9.4.1 Defining an instrument method	114
2.9.4.2 Acquiring data using the static nanospray	116
2.10 Protein analysis using BioWorks v 3.2 software (ThermoElectron)	116
2.10.1 Creating a streptococci database	116
2.10.2 Performing TurboSEQUENT searches	117
2.10.3 Filtering the match list.....	118
2.11 Protein identification and characterisation using TSSView™ v1.0 (Turbinia/Biosystems Informatics Institute)	119
2.11.1 Protein identification via functional conserved sequence.....	119
2.11.2 Functional characterisation of identified proteins	119
3 RESULTS FOR DEVELOPMENT OF TECHNOLOGY FOR PROTEOMICS USING 2DE WITH MS.....	120
3.1 Background.....	120
3.2 Development of 2DE	121
3.2.1 2DE of <i>E. coli</i> TOP10	121
3.2.1.1 Extraction of cell-associated proteins from <i>E. coli</i> TOP10	121
3.2.1.2 Experiments to optimize conditions for basic 2DE	121
3.2.2 2DE of streptococci.....	128
3.2.2.1 Development of 2DE for separation of proteins from <i>S. pyogenes</i> SF370	128
3.2.2.1.1 Separation of cell-associated proteins in pH range 3 -10	128
3.2.2.1.2 Separation of cell-associated proteins in pH range 4 – 7	135
3.2.2.1.3 Separation of secreted proteins from culture supernatant in pH range 3 – 10	140
3.3 Development of static nanospray MS analysis and protein identification using BioWorks v 3.2 software.....	142
3.3.1 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS.....	142
3.3.2 Purified protein identification using static nanospray LCQ Advantage MS with BioWorks 3.2 software.....	142

4 RESULTS FOR PROTEOMIC ANALYSIS OF STREPTOCOCCI GROWN IN HA-RICH MEDIA.....	152
4.1 Background.....	152
4.2 Proteomic analysis of GAS grown in HA-rich media	153
4.2.1 Growth of streptococci in defined media	153
4.2.2 Growth of <i>S. pyogenes</i> SF370 in HA-rich media and control media	153
4.2.3 Investigation into the expression of hyaluronate lyases in <i>S. pyogenes</i> SF370 cultures	155
4.2.4 Investigation into the consumption of carbohydrates by <i>S. pyogenes</i> SF370	157
4.2.5 Comparison of differentially expressed cell-associated proteins of <i>S. pyogenes</i> SF370 growing in HA-rich and control media using 2DE..	157
4.2.6 Analysis of protein expression of GAS growing in HA-rich media and control media using PDQuest™ v 6.2.1 software	159
4.2.7 Identification of GAS proteins from 2DE gels using static nanospray LCQ Advantage MS with BioWorks 3.2 software	168
4.2.8 Protein identification and characterisation using TSSView™ v 1.0	168
4.2.8.1 Protein identification via highly-conserved sequence identification	168
4.2.8.2 Functional analysis of hypothetical proteins	186
4.2.8.3 Functional characterisation of identified proteins	189
4.2.9 Comparison of differentially expressed secreted proteins of <i>S. pyogenes</i> SF370 growing in HA-rich media and control media using 2DE	189
5 DISCUSSIONS	194
5.1 Development of techniques for proteomics.....	194
5.1.1 Development of techniques for 2DE	194
5.1.2 Analysis of 2DE gels using PDQuest™ software.....	197
5.1.3 Development of techniques for static nanospray LCQ Advantage MS.....	197
5.1.4 Protein identification.....	199
5.1.5 Functional analysis of proteins.....	200
5.2 Proteomics analysis of <i>S.pyogenes</i> SF370 growing in HA-rich media....	202
5.2.1 Growth of <i>S. pyogenes</i> SF370 in HA-rich media and control media	203
5.2.2 Analysis of expression of hyaluronate lyase by <i>S. pyogenes</i> SF370	204
5.2.3 Analysis of differentially expressed proteins in HA-rich media ..	205
5.3 Conclusion.....	210
5.4 Future work	212
6 REFERENCES	214
APPENDICES	246
Appendix A	246
A1: Chemicals, media and enzymes used in this study.....	246

A2: Protein size standards	249
A3: 2-D Clean-Up kit	249
A4: 2-D quant kit.....	250
Appendix B	252
B1: Autoclaving	252
B2: pH meter.....	252
B3: Spectrophotometer	252
B4: Incubators	252
B5: Centrifugation.....	252
B6: Sonication	253
B7: Microtitre plate reader	253
B8: SDS-PAGE gel kit.....	253
B9: Gel documentation	253
B10: Immobiline DryStrip Kit	254
B11: Protean II XL 2-D cells and accessories	254
B12: Centrifugal evaporator	256
Appendix C Preparation of dialysed sodium hyaluronate	257
Appendix D	259
Appendix E	264
Appendix F.....	271
Appendix G	319
PUBLISHED PAPER.....	333

FIGURES

FIGURE 1.1: NORMAL CONNECTIVE TISSUE.....	19
FIGURE 1.2: DISACCHARIDE REPEATING UNIT OF HA COMPRISING <i>N</i> -ACETYLGLUCOSAMINE AND GLUCURONIC ACID.....	22
FIGURE 1.3: CHAIN OF GROUP A STREPTOCOCCUS.....	24
FIGURE 1.4: CELL SURFACE STRUCTURE OF <i>S. PYOGENES</i> AND SECRETED PRODUCTS INVOLVED IN VIRULENCE.....	28
FIGURE 1.5: STRUCTURE OF M PROTEINS.....	31
FIGURE 1.6: MUCOID COLONIES OF <i>S. PYOGENES</i> ON BLOOD AGAR.....	33
FIGURE 1.7: COMPLETELY SEQUENCED GENOMES ACCORDING TO GENOMESONLINE.....	44
FIGURE 1.8: GENOME SEQUENCING PROJECTS ACCORDING TO GENOMESONLINE.....	45
FIGURE 1.9: THE PERCENTAGE OF GENES IN EACH ROLE ACROSS ALL PUBLISHED GENOMES.....	51
FIGURE 2.1: THE B-ELIMINATION REACTION CATALYZED BY A HYALURONATE LYASE.....	70
FIGURE 2.2: APPLYING SAMPLE INTO SAMPLE CUPS.....	82
FIGURE 2.3: PROCESS INVOLVED IN ANALYSIS OF 2DE USING PDQUEST™ V 6.2.1 SOFTWARE.....	99
FIGURE 2.4: STATIC NSI PROBE ASSEMBLY AND STATIC NSI SOURCE.....	109
FIGURE 3.1: VISUALIZATION OF CELL-ASSOCIATED PROTEINS FORM <i>E. COLI</i> ON 12 % (W/V) SDS-PAGE.....	122
FIGURE 3.2: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>E. COLI</i> (EXPERIMENTS NO.1).....	124
FIGURE 3.3: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>E. COLI</i> (EXPERIMENTS NO.2).....	125
FIGURE 3.4: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>E. COLI</i> (EXPERIMENTS NO.1).....	126
FIGURE 3.5: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>E. COLI</i> (EXPERIMENTS NO.1).....	127
FIGURE 3.6: 2DE GEL OF CELL-ASSOCIATED PROTEINS FORM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.5).....	130
FIGURE 3.7: 2DE GEL OF CELL-ASSOCIATED PROTEINS FORM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.6).....	131

FIGURE 3.8: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.7).....	132
FIGURE 3.9: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.8).....	133
FIGURE 3.10: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.9).....	134
FIGURE 3.11: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.10).....	137
FIGURE 3.12: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.11).....	138
FIGURE 3.13: 2DE GEL OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370 (EXPERIMENT NO.12).....	139
FIGURE 3.14: 2DE GEL OF SECRETED PROTEINS FROM <i>S. PYOGENES</i> SF370 CULTURE SUPERNATANT (EXPERIMENT NO.13).....	141
FIGURE 3.15: 12 % (W/V) SDS – PAGE ANALYSIS OF DIGESTION OF PROTEINS.....	143
FIGURE 3.16: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING ALBUMIN DIGESTS.....	144
FIGURE 3.17: 12 % (W/V) SDS – PAGE ANALYSIS OF UGL DIGESTION.....	150
FIGURE 3.18: MASS SPECTRUM OF UGL DIGESTION.....	150
FIGURE 4.1: GROWTH CURVES FOR <i>S. PYOGENES</i> SF370 IN HA-RICH MEDIA AND BHI MEDIA.....	154
FIGURE 4.2: THE ABSORBANCE AT 232 NM FOR INVESTIGATION INTO THE EXPRESSION OF HYALURONATE LYASES.....	158
FIGURE 4.3: COMPARISON OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370.....	160
FIGURE 4.4: 2DE GEL ANALYSIS OF CELL-ASSOCIATED PROTEINS FROM <i>S. PYOGENES</i> SF370.....	169
FIGURE 4.5: IDENTIFICATION OF PROTEIN SPOT NO.1 AS A PUTATIVE PHOSPHOTRANSFERASE SYSTEM PHOSPHOHISTIDINE-CONTAINING PROTEIN.....	171
FIGURE 4.6: IDENTIFICATION OF PROTEIN SPOT NO.2 AS A PUTATIVE NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.....	172
FIGURE 4.7: IDENTIFICATION OF PROTEIN SPOT NO.3 AS A M1 PROTEIN PRECURSOR.....	173
FIGURE 4.8: IDENTIFICATION OF PROTEIN SPOT NO.4 AS A HYPOTHETICAL PROTEIN SPY1143.....	174

FIGURE 4.9: IDENTIFICATION OF PROTEIN SPOT NO.5 AS A HISTIDINE PROTEIN KINASE.....	175
FIGURE 4.10: IDENTIFICATION OF PROTEIN SPOT NO.6 AS A TRNA MODIFICATION GTPASE.....	176
FIGURE 4.11: IDENTIFICATION OF PROTEIN SPOT NO.7 AS A HYPOTHETICAL PROTEIN SPY1262.....	177
FIGURE 4.12: IDENTIFICATION OF PROTEIN SPOT NO.8 AS A 50S RIBOSOMAL PROTEIN L10.....	178
FIGURE 4.13: IDENTIFICATION OF PROTEIN SPOT NO.9 AS A ROPA.....	179
FIGURE 4.14: IDENTIFICATION OF PROTEIN SPOT NO.10 AS A COLLAGEN-LIKE SURFACE PROTEIN.....	180
FIGURE 4.15: IDENTIFICATION OF PROTEIN SPOT NO.11 AS A PUTATIVE CARBOMOYL-PHOSPHATES SYNTHASE SMALL SUBUNIT.....	181
FIGURE 4.16: IDENTIFICATION OF PROTEIN SPOT NO.12 USING TURBOSEQUENT.....	182
FIGURE 4.17: IDENTIFICATION OF PROTEIN SPOT NO.13 USING TURBOSEQUENT.....	183
FIGURE 4.18: IDENTIFICATION OF PROTEIN SPOT NO.15 USING TURBOSEQUENT.....	184
FIGURE 4.19: IDENTIFICATION OF PROTEIN SPOT NO.18 USING TURBOSEQUENT.....	185
FIGURE 4.20: IDENTITY OF REGIONS OF HIGHLY-CONSERVED REGIONS IN THE HYPOTHETICAL PROTEIN SPY1143.....	188
FIGURE 4.21: IDENTITY OF REGIONS OF HIGHLY-CONSERVED REGIONS IN THE HYPOTHETICAL PROTEIN SPY1262.....	188
FIGURE 4.22: IDENTITY OF REGIONS OF HIGHLY-CONSERVED REGIONS IN M1 PROTEIN.....	190
FIGURE 4.23: IDENTITY OF REGIONS OF HIGHLY-CONSERVED REGIONS IN TRNA MODIFICATION GTPASE.....	191
FIGURE 4.24: COMPARISON OF SECRETED PROTEINS FROM <i>S.</i> <i>PYOGENES</i> SF370 CULTURE SUPERNATANT IN PH RANGE 3-10.....	192
FIGURE 4.25: COMPARISON OF SECRETED PROTEINS FROM <i>S.</i> <i>PYOGENES</i> SF370 CULTURE SUPERNATANT IN PH RANGE 6-11.....	193
FIGURE 5.1: A VIEW OF BIOINFORMATICS PAST, PRESENT AND FUTURE.....	201
FIGURE B1 IMMOBILINE DRYSTRIP KIT.....	254
FIGURE B2: PROTEAN II XL 2-D CELLS.....	255
FIGURE B3: P ROTEAN II XL 2-D ACCESSORIES.....	255

FIGURE D1: TYPICAL STANDARD CURVE FOR DNSA ASSAYS.....	259
FIGURE D2: STANDARD CURVE FOR DNSA REDUCING SUGAR ASSAYS.....	260
FIGURE D3: STANDARD CURVE FOR THE URONIC ACID ASSAY.....	261
FIGURE D4: BRADFORD ASSAY STANDARD CURVE FOR THE QUANTIFICATION OF PROTEIN.....	262
FIGURE D5: TYPICAL STANDARD CURVE FOR ASSAY USING 2-D QUANT KIT.....	263
FIGURE E1: THE REPLICATES OF THE 2D GELS IN EXPERIMENT NO.5 – 9.....	264
FIGURE E2: THE REPLICATES OF THE 2D GELS IN EXPERIMENT NO.10 – 11.....	269
FIGURE F1: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING A- LACTALBUMIN DIGESTS.....	271
FIGURE F2: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING B-GALACTOSIDASE DIGESTS.....	277
FIGURE F3: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING PHOSPHORYLASE B DIGESTS.....	283
FIGURE F4: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING L-GLUTAMIC DEHYDROGENASE DIGESTS.....	289
FIGURE F5: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING OVALBUMIN DIGESTS.....	295
FIGURE F6: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE DIGESTS.....	301
FIGURE F7: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING CARBONIC ANHYDRASE DIGESTS.....	307
FIGURE F8: ANALYSIS OF SALT AND DETERGENT TOLERANCE OF THE STATIC NANOSPRAY LCQ ADVANTAGE MS USING TRYPSIN DIGESTS.....	313
FIGURE G1: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM PUTATIVE PHOSPHOTRANSFERASE SYSTEM PHOSPHOHISTIDINE-CONTAINING PROTEIN.....	319
FIGURE G2: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM PUTATIVE NADP-DEPENDENT GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE.....	320

FIGURE G3: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM M1 PROTEIN PRECURSOR.....	321
FIGURE G4: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM HYPOTHETICAL PROTEIN SPY1143.....	322
FIGURE G5: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM HISTIDINE PROTEIN KINASE.....	323
FIGURE G6: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM TRNA MODIFICATION GTPASE.....	324
FIGURE G7: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM HYPOTHETICAL PROTEIN SPY1262.....	325
FIGURE G8: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM50S RIBOSOMAL PROTEIN L10.....	328
FIGURE G9: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM ROPA.....	329
FIGURE G10: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM COLLAGEN-LIKE SURFACE PROTEIN.....	331
FIGURE G11: MS/MS SPECTRA OF IDENTIFIED PEPTIDES FROM PUTATIVE CARBOMOYL-PHOSPHATES SYNTHASE SMALL SUBUNIT.....	332

TABLES

TABLE 2.1: THREE DEFINED MEDIA WITH SERIAL CONCENTRATION OF GLUCOSE.....	59
TABLE 2.2: THREE DEFINED MEDIA WITH SERIAL CONCENTRATION OF HA.....	61
TABLE 2.3: SOURCES USED IN THE DETECTION OF THE EXPRESSION OF HYALURONATE LYASE.....	66
TABLE 2.4: EIGHT SAMPLES WERE USED IN DNSA ASSAY.....	68
TABLE 2.5: FOUR REACTIONS WERE SET UP FOR ANALYSIS OF CONSUMPTION OF CARBOHYDRATES IN <i>S. PYOGENES</i> SF370 CULTURES.....	72
TABLE 2.6: RUNNING CONDITIONS FOR IMMOBILINE DRYSTRIP PH 3-10 USING MULTIPHOR II UNIT.....	84
TABLE 2.7: PREPARATION OF STANDARD CURVE FOR ASSAY USING 2-D QUANT KIT.....	94
TABLE 2.8: RUNNING CONDITIONS FOR IMMOBILINE DRYSTRIP PH 4-7 USING MULTIPHOR II UNIT.....	96
TABLE 2.9: RUNNING CONDITIONS FOR IMMOBILINE DRYSTRIP PH 6-11 USING MULTIPHOR II UNIT.....	96
TABLE 2.10: SAMPLE FOR ANALYSIS OF SALTS AND DETERGENT TOLERANCE.....	113
TABLE 3.1: INFORMATION OF MATCHED PEPTIDES FROM UGL.....	151
TABLE 3.2: PROTEIN COVERAGE OF THE MATCHED PEPTIDES OF UGL.....	151
TABLE 4.1: GROWTH OF <i>S. PYOGENES</i> SF370 IN DEFINED MEDIUM....	154
TALBE 4.2: THE CONCENTRATIONS OF REDUCING SUGAR IN DIFFERENT SAMPLES FOR INVESTIGATING THE EXPRESSION OF HYALURONATE LYASES.....	156
TABLE 4.3: BLASTPAM30 ALIGNMENTS FOR TWO PEPTIDES OF PUTATIVE THE FORMATETETRAHYDROFOLATE LIGASE.....	187
TABLE A1: MOLECULAR WEIGHT DISTRIBUTION IN SIGMA MARKERS.....	249
TABLE B1: CENTRIFUGE ROTORS, VESSELS AND APPLICATIONS.....	253

Acknowledgements

I wish to thank my supervisor Prof. Gary Black for his immeasurable contribution to this work over the past years, Prof. Ian Humphery-Smith for his advice and guidance, Dr. Simon Charnock for his help at the beginning of the project, and Dr John Perry for providing bacterial lenticules. I would also like to thank my Mother for all she is, and all she has done for me. I acknowledge the efforts of Miss. Fiona McDonald and Dr Shane Sturrock for their help in Bioinformatics analysis. I appreciate the help from Dr. Nicola Smith, Dr Joanne Baxter, Miss Anna, Lindsay, Ms. Zainab Mabrounk, Miss Mandy Lyall, Mr. Bill Sheldon and members of A307 past and present. Also thank to Universities UK for project funding.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work.

Name:

Meng Zhang

Signature:

Date:

Abbreviations

ACN	Acetonitrile
APS	Ammonium persulphate
BHI	Brain heart infusion
BHI + HA	BHI with 0.5 % HA
BPB	Bromophenol blue
BSA	Bovine serum albumin
°C	Degree Celsius
C-terminal	Carboxy terminal
Da	Dalton
2DE	Two dimensional gel electrophoresis
DNA	Deoxyribonucleic acid
DNSA	Dinitrosalicylic acid
DTT	Dithiothreitol
ESI	Electrospray ionization
g	Gram(s)
GAS	Group A streptococcus
h	Hour(s)
HA	Hyaluronan/Hyaluronic acid
HMW	Low molecular weight
IAA	Equilibration with iodoacetamide
LMW	Low molecular weight
Ig A	Immunoglobulin A
Ig G	Immunoglobulin G
kb	Kilobase pair(s)
kDa	Kilodalton
L	Litre(s)
LB	Luria-Bertani medium
LTA	Lipoteichoic acid
m	Metre(s)
M	Molar / Methionine
mA	Milliamps

min	Minute(s)
mm	Millimetre(s)
M_r	Relative molecular mass
MS	Mass spectrometry
N-terminal	Amino terminal
OD _x	Optical density at x nm
PAM	Plasminogen
PBS	Phosphate Buffered Saline
PMNs	Polymorphonuclear leukocytes
rpm	Revolutions per minute
s	Second(s)
SDS	Sodium dodecylsulphate
SLO	Streptolysin O
SLS	Streptolysin S
Spe	Streptococcal pyrogenic exotoxins
SpeB	Cysteine protease
STSS	Streptococcal toxic shock syndrome
SLO	Streptolysin O
SLS	Streptolysin S
TEMED	N,N,N',N'-tetramethylethylene diamine
TFA	Trifluoroacetic acid
TIGR	The Institute for Genomic Research
Tris	tris(hydroxymethyl)aminomethane
UV	Ultraviolet
V	Volt(s)
v/v	Volume per volume
w/v	Weight per volume
x g	Times gravity
α	Alpha
β	Beta
Δ	Delta
k	Kilo
λ	Lambda

m	milli
μ	Micro
n	Nano
18.2 MΩ H ₂ O	18.2 mega ohm water

1 Introduction

1.1 Connective tissues

Connective tissues are ubiquitous in the mammalian body. Thibodeau and Patton suggested that connective tissues are the most abundant and widely distributed tissues in the body (Thibodeau and Patton, 1992). It is found in almost every organ, including skin, membranes, muscles, bones and nerves (Johnson, 1991). Connective tissues have a number of different functions depending on their structure and appearance. The basic function of connective tissues is, as the name implies, connecting. They provide a structural component; a form a packing material to fill up the unoccupied space and hold and support other tissues and organs in the body. Connective tissues also provide metabolic support. They can deposit fat and mediate the exchange of nutrients, metabolites and waste products between tissues and organs. Due to their structure, connective tissues can also play an important role in the body's defences against pathogenic microorganisms (Burkitt, 1993). Therefore, during the infection of pathogens, the degradation of connective tissue by pathogens is a key element.

1.1.1 General connective tissues

All connective tissues are made up of cells and intercellular substances. The intercellular substances consist of extracellular fibres and an amorphous ground substance, which is also called 'matrix' (Figure 1.1). The fibrous components of supporting tissue are of two main types, collagen fibres, including reticular fibres, and elastic fibres. Collagen fibres are the main fibre type and the most abundant protein in the human body. The most notable function is the provision of tensile strength. The elastic fibres are composed of elastin, and they exist in skin, lung and blood vessels where they confer the properties of elastic recoil to enable recovery of tissue shape following normal physiological deformation (Burkitt, 1993). According to whether its constituent fibres are loosely or tightly packed, connective tissues are classified into loose connective tissue and dense connective tissue. Loose connective tissues include areolar connective tissue, which is widely distributed in the body and form the fascia

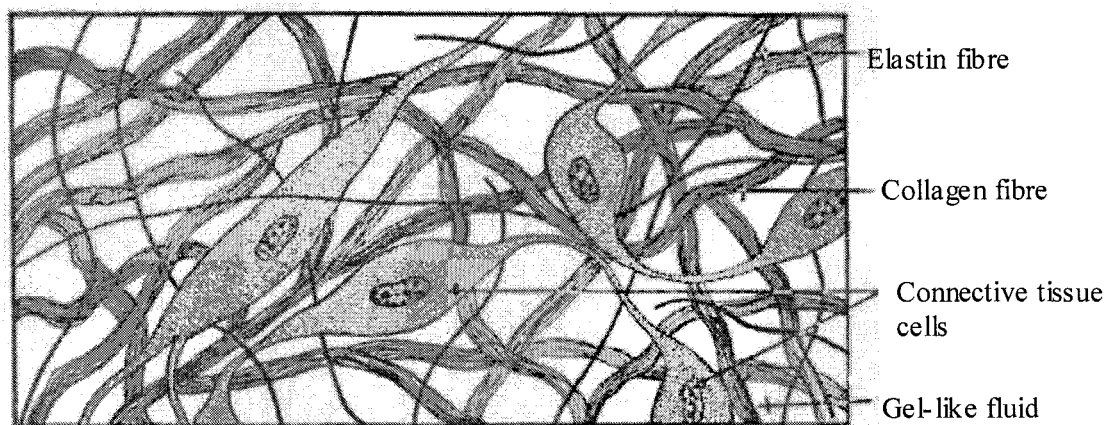


Figure 1.1 Normal connective tissue.

Connective tissue forms the glue between cells in the body's tissues, holding the cells together and giving those tissues structure and shape. It is made of a number of different molecules, including a range of fibres that give strength and elasticity to many tissues.

<http://www.marfan.net.au/>.

that binds the organs together (Krause, 1996). Adipose tissue is another kind of loose connective tissue, and it is a specialized connective tissue that is a major storage site for fat, in the form of triglycerides. Adipose tissue fulfils several functions including storing fat and assisting thermoregulation (Johnson, 1991). Both loose connective tissues and dense connective tissues are general connective tissues.

1.1.2 Special connective tissues

Special connective tissues have unique histological organization and special functions which make them distinct from other general connective tissues. Cartilage, bone and blood are special connective tissues. Identical to general connective tissues, cartilage and bone are composed of cells, extracellular fibres and amorphous ground substance. The ground substance of cartilage, which has high water content, results in firm consistency and ability to maintain the tissue shape. Closely packed collagenous fibres embedded into the ground substance provide tensile strength and elasticity. The special structural organization of cartilage matrix gives it the ability to absorb and distribute loads (Huber et al., 2000). In addition, articular cartilage that covers the ends of long bones in synovial joints differs from other connective tissues in that it contains no separating basement membranes, nerves, lymphatics or blood vessels, and the access to nutrients and elimination of waste products occur via diffusion through the extracellular matrix (Kuettner, 1992). Bone is a type of mineralised connective tissue and in addition to collagen fibres it contains a large store of calcium phosphate which gives bone its firmness. Bones are the principle supports of the body and they provide a rigid structure for muscle attachment with locomotion and cartilage augments this structure by providing flexible support. Cartilage and bone have many common features, and in endochondral ossification, cartilage models are converted to bone (Johnson, 1991).

Although blood is a fluid, it is classified as connective tissue as it consists of a large amount of ground substance, blood plasma, and blood cells including erythrocytes, leukocytes and platelets. However, unlike other connective tissues, the intercellular substance of blood lacks a fibrous component. Blood is important in the transport of gases, nutrients, metabolic waste products, cells and hormones throughout the body

and it also aids in regulating body temperature by dissipating heat formed during metabolism (Burkitt, 1993).

1.1.3 Amorphous ground substance

In connective tissues, cells and fibres are embedded in the amorphous ground substance, which is a transparent semi-fluid gel of variable viscosity. Ground substance consists of glycoproteins, glycosaminoglycans and proteoglycans. Glycosaminoglycans are long unbranched polysaccharides of repeating disaccharide units. The disaccharide units are usually a uronic acid, such as glucuronate or iduronate, and either of two modified sugars, *N*-acetylgalactosamine or *N*-acetylglucosamine (Burkitt, 1993). Due to the presence of hydroxyl, carboxyl and sulphate side groups on the disaccharide units, glycosaminoglycans are highly negatively charged molecules. Hyaluronic acid (HA) is the chief glycosaminoglycan of loose connective tissue, and is the only one without sulphate side groups. HA is a high molecular mass polysaccharide (relative molecular mass of $\sim 1,000,000$) consisting of highly repeated disaccharide units, approximately 2,000 to 13,000 times, of *N*-acetylglucosamine and glucuronic acid (Figure 1.2). HA is distributed in various tissues, such as liver, blood, brain, the aorta, skin and the vitreous humor of the eye (Lindahl and Hook, 1978), with the highest concentrations found in soft connective tissue (Laurent and Fraser, 1992). It is synthesized in the plasma membrane, unlike the other glycosaminoglycans, including chondroitin-4-sulphate, chondroitin-6-sulphate, dermatan sulphate and heparin sulphate, which are made in the Golgi. These other glycosaminoglycans differ from HA in that they are covalently linked to a variety of core protein molecules to form proteoglycans, which are huge molecules consisting of 90-95% carbohydrate (Burkitt, 1993). Further, HA links to the proteoglycans with non-covalent bound to form even larger molecular complexes. In addition, ground substance contains large amount of water, which is bound to the glycosaminoglycans and proteoglycans. The ability of HA to bind water is responsible for the viscosity of loose connective tissue. The size of the space between these large molecular complexes is normally less than 0.5×10^{-6} m, a typical bacterium size, and therefore determines the permeability of the connective tissues and aids in inhibiting or retarding the spread of pathogens and their toxic materials from the site of an infection (Krause, 1996).

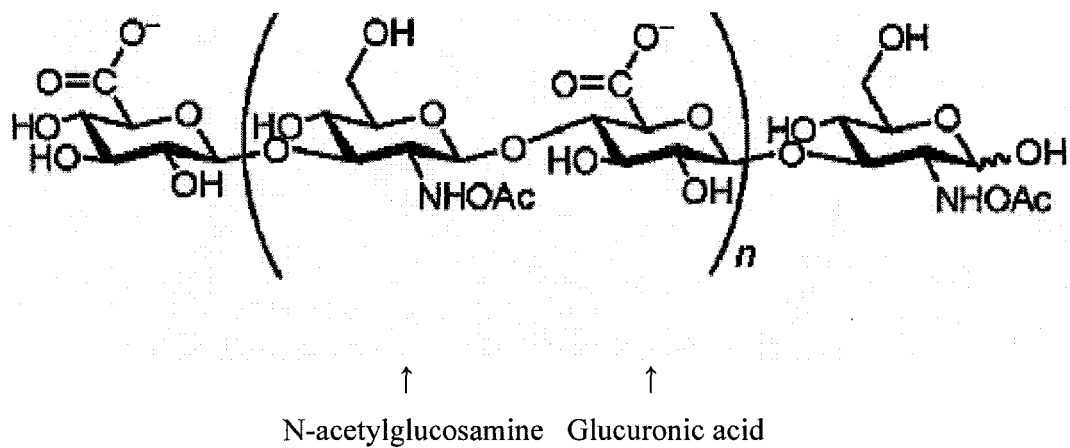


Figure 1.2 Disaccharide repeating unit of HA comprising *N*-acetylglucosamine and glucuronic acid

http://www.checue.uq.edu.au/research/bioengineering/research/Metabolic_Engineering/HA.html

1.2 Streptococci

Streptococci are Gram-positive bacteria. Various streptococci are important ecologically as one of the normal microbial flora of the host, such as humans, horses, cows and pigs. Some can also cause diseases that range from subacute to acute, or even chronic. When streptococci infect humans, the infections normally occur in the respiratory tract, the skin, wounds, the bloodstream and muscle tissues (Salyers and Whitt, 2002). Human diseases are most commonly associated with *Streptococcus pyogenes*. As the 'prefix' 'Strep' indicates, streptococci exist as chains, where the bacteria undergo division along the same plane and tend to remain together without separating after cell division (Figure 1.3)

1.2.1 Classification of streptococci

1.2.1.1 Haemolysis

Streptococci growing on blood agar plates secrete a variety of haemolysins. Some can completely lyse nearby blood cells to create a zone of clearing around colonies, and this type of hemolysis is called β -hemolysis. Less complete lysis of blood cells, causing a green zone to form, is called α -hemolysis, and γ -hemolysis does not produce a zone at all (Greenwood et al., 1992). *S. pyogenes* causes β -hemolysis.

1.2.1.2 Lancefield grouping

Streptococci are also classified into several groups, based on the immunological differences in their cell wall polysaccharides, as is the case with Group A, B, C, F and G streptococcus, or their lipoteichoic acid, as is the case with Group D streptococcus. This classification was proposed by Dr. Rebecca Lancefield and therefore the groups are known as the Lancefield groups (Greenwood et al., 1992). *S. pyogenes* is group A streptococcus (GAS).

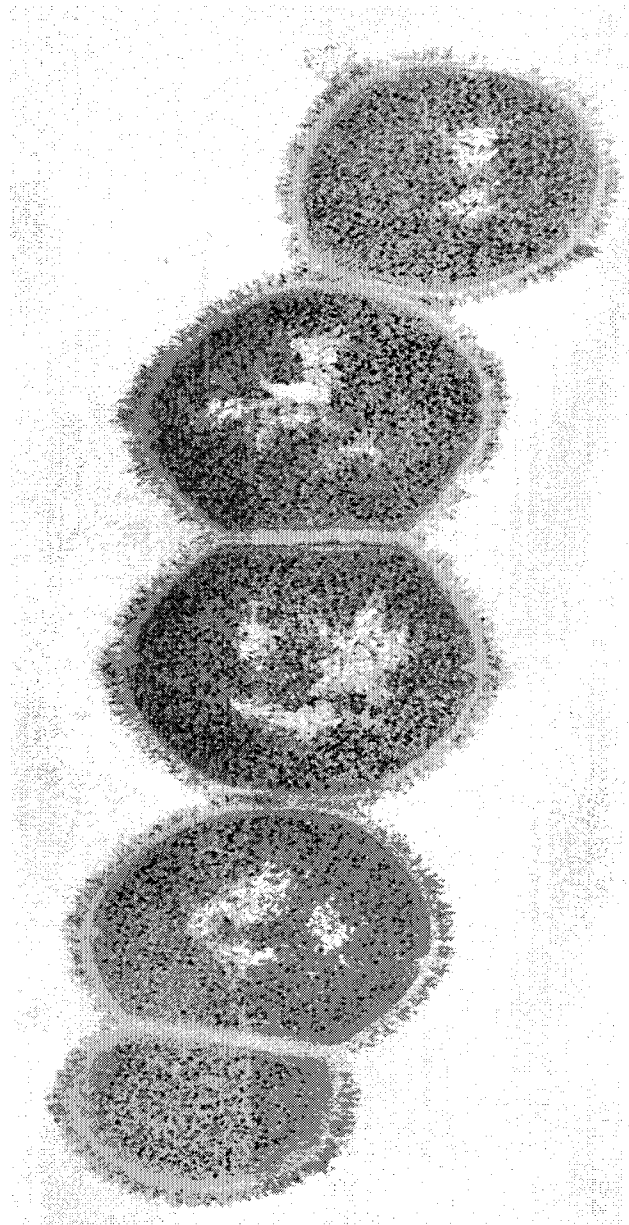


Figure 1.3 Chain of Group A streptococcus.

Streptococci exist as chains, where these bacteria undergo division along the same plane and tend to remain together without separating after cell division.

<http://www.rockefeller.edu/vaf/chain20.htm>

In addition, Lancefield also showed the further subdivision of GAS into specific serological types based on the presence of type specific M (protein) antigens (Rotta et al., 1971). The M protein is a main surface protein of GAS, and it can be extracted from the bacteria with boiling hydrochloric acid (Cunningham, 2000). M proteins are very antigenic, and after the host develops an antibody response against them, the bacteria are readily recognized and destroyed. Thus, there is selective pressure to alter the structure of M proteins. According to Bisno's review, there are 124 recognised M genotypes, and they are usually called M1 GAS, M2 GAS, M3 GAS.....etc. It was assumed that M proteins are strain specific (Johnson and Kaplan, 1993), however in 1994, Whatmore and Kehoe provided the first evidence that horizontal transfer of M-like sequence occurred between distinct GAS strains which suggested that a single M protein type can include very different *S. pyogenes* strains (Whatmore and Kehoe, 1994). Also due to alternations occurring during DNA replication (Bisno, 2003), more than one strain can exist in any one M serotype.

1.2.2 GAS and GAS disease

GAS is a human pathogen. It is estimated that between 5 - 15 % of normal individuals harbour the bacteria, usually in the respiratory tract, without signs of diseases. However, as normal flora, GAS can cause infection when host defenses are compromised or when the organisms are able to penetrate the constitutive host defenses. When the bacteria are able to penetrate the vulnerable tissue, a variety of types of suppurative infection can occur.

The primary locations of GAS infections are the throat, the skin, and even in deeper sites like the blood, muscle and the lungs. GAS is the most common bacterial cause of strep throat, also called streptococcal sore throat, which includes pharyngitis, the infection of the mucous membranes lining the pharynx, and tonsillitis, the infection of the tonsil. Infections produce symptoms including high fever chills, headache, muscle aches, bright red tongue, swollen lymph glands and nausea. Although anyone can get strep throat, it primarily affects school-age children, and it has been concluded that approximately 15 % of school-age children suffer pharyngitis each year in developed countries (Carapetis et al., 2005). GAS can spread to adjacent tissues to produce otitis

media and sinusitis. When GAS invades the blood it can cause septicaemia and even meningitis sometimes if it also invades the cerebrospinal fluid (van de Beek et al., 2002).

GAS may cause several types of skin infections, including impetigo or pyoderma, which mainly affects exposed areas on children's faces, arms and leg and can spread between hosts. Erysipelas is a more severe streptococcal skin infection. It occurs in the superficial layers of the skin (cellulites) and involves the lymphatics. Erysipelas is manifested by local signs of inflammation (warmth, erythema, and pain) and, in most instances, by fever and leukocytosis (Bisno and Stevens, 1996). Skin infections usually occur during the summer months and in temperate climates. GAS can also cause wound infections. Necrotizing fasciitis is a type of wound infection, which is caused by the entry of GAS into the soft tissues through the skin. This infection progresses very rapidly, destroying fat and fascia (Bisno and Stevens, 1996, Greenwood et al., 1992). Patients with necrotizing fasciitis may develop streptococcal toxic shock syndrome (STSS), which is caused by streptococcal toxins released by GAS into the blood stream. The early symptoms of STSS include myalgia, malaise, chills, fever, nausea, vomiting, and diarrhea. In some patients, particularly those with necrotizing fasciitis, pain at the site of minor trauma may be the initial symptom. In addition, due to the release of streptococcal pyrogenic exotoxins (SPE), GAS can cause scarlet fever. In scarlet fever, a hypersensitivity reaction to these toxins gives rise to a red rash. Another type of invasive infection is myositis, which is occurred by bacteria entering into the muscle.

Immune-mediated post-streptococcal sequelae, such as acute rheumatic fever and acute glomerulonephritis may also be developed following acute infections caused by GAS (Olivier, 2000, Cunningham, 2000). Acute glomerulonephritis and acute rheumatic fever occur at least as commonly in adults as children. Acute glomerulonephritis is triggered by GAS skin infection, and acute rheumatic fever occurs only after pharyngeal infections. Presenting signs and symptoms of glomerulonephritis include uraemia, oedema and hypertension (Stulberg et al., 2002). Symptoms of acute rheumatic fever include polyarticular arthritis, carditis, chorea, erythema marginatum and/or subcutaneous nodules (Greenwood et al., 1992). Rheumatic fever can progress to damage the heart valves (rheumatic heart disease), and it is a major cause of

acquired heart disease in young people throughout the world. Fortunately, rheumatic fever and rheumatic heart disease have largely disappeared from developed countries (Salyers and Whitt, 2002).

It has been reported that GAS infections globally affect more than 18 million people per year, with more than 0.5 million deaths per year (Carapetis et al., 2005). Therefore, GAS infections remain a worldwide problem, and further research is needed to reveal the mechanism of pathogenesis.

1.2.3 GAS virulence factors

It is well known that many diseases caused by GAS are the results of the invasion of GAS in the host tissues by producing a wide range of virulence factors. In the past, intensive investigation has been done to study these virulence factors, and the well known virulence factors include M proteins, M-like proteins, the HA capsule, C5a peptides, lipoteichoic acid, fibronectin binding proteins, cysteine protease (SpeB), hyaluronate lyase, streptokinase, DNAases, streptolysins and superantigenic exotoxins (Cunningham, 2000, Hynes, 2004, Bisno et al., 2003, Greenwood et al., 1992, Mims et al., 1998) (Figure 1.4).

1.2.3.1 M proteins and M-like proteins

When GAS invades the host tissue, the key component of host defence against infection is the recruitment of polymorphonuclear leukocytes (PMNs) to ingest and kill GAS (Bisno et al., 2003). To survive in the host tissue, GAS produces several virulence factors to avoid the recognition by PMNs, including M proteins.

M proteins are fibrous rodlike molecules that are α -helical homodimers arranged in a coiled-coil form. Basically, there are two α -helices that coil around one each other. The amino-terminal region of an M protein consists of three distinct blocks of tandem repetitive sequences referred as the A, B and C repeat domains, and the carboxyl-terminal region contains a LPXTG motif, which covalently links the M protein to the

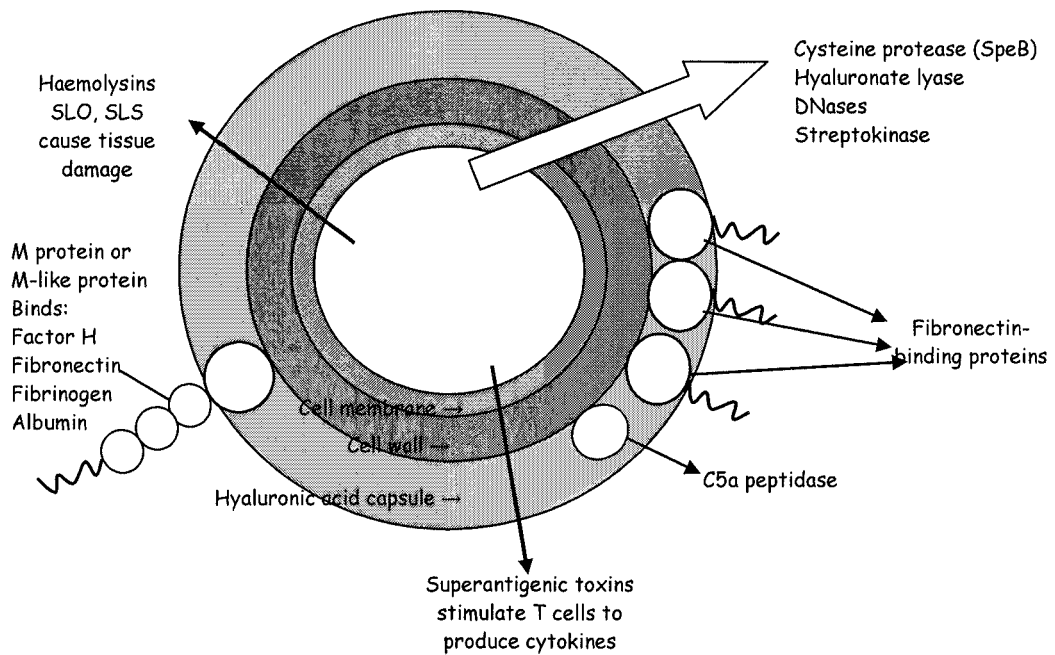


Figure 1.4 Cell surface structure of GAS and secreted products involved in virulence.

The cell envelope of a GAS is illustrated. GAS produces a wide range of virulence factors to allow adhesion and evasion. M protein, M-like proteins and multiple fibronectin binding proteins contribute to GAS adhesion. M protein, HA capsule and C5a peptidase play roles in anti-phagocytosis. Spe B, SLS, SLO and DNase cause cell and tissue damage. Superantigenic toxins, Streptokinase, SLS, SLO and DNase play roles in evasion of host immune response.

cell wall peptidoglycan (Figure 1.5) (Caparon, 2001, Husmann et al., 1997, Fischetti, 1989). Although the sequence and size of M proteins differ from strain to strain, this basic structure is conserved.

M proteins are major surface proteins that are defined by their ability to protect the GAS from phagocytosis by PMNs. Many studies have shown that M-negative GAS is effectively opsonized and as a result, GAS are rapidly ingested and killed by PMNs (Peterson et al., 1979, Bisno, 1979). In contrast, GAS rich in surface M proteins are neither opsonized nor ingested (Lancefield, 1962). Furthermore, when type-specific M protein antibody was added to human blood, M rich GAS were opsonized and easily ingested and killed by PMNs (Stollerman et al., 1958). In the last 50 years, many studies have been performed to explain the antiphagocytic properties of streptococcal M proteins. In 1979, Bisno suggested that M proteins can prevent opsonization by interfering with the activation of an alternative complement pathway (Bisno, 1979). The deposition of component C3b on GAS surface allows opsonization. However, M proteins are able to bind complement factor H, which is a 150 kDa 3-globulin present in serum and controls the destruction of surface-bound C3b. This activity discourages the deposition of C3b on GAS surface and results in the prevention of effective opsonization of GAS (Fischetti, 1989, Salyers and Whitt, 2002). Then, Kotarsky et al reported that M proteins bind to factor H via the C-repeat region (Kotarsky et al., 2001). It also has been showed that M protein binds to a C4b binding protein and factor H-like protein at the hypervariable N-terminal region. C4b binding protein is a complement regulator and promotes decay of the C3 convertase from the classical pathway of human complement and acts as a cofactor to the plasma protease factor I in the degradation of C4b. Factor H-like protein has a function similar to C4b binding protein, but it is involved in the inhibition of the C3 convertase from the alternative pathway of human complement and in the inactivation of C3b. M protein binding to H-like protein and C4b binding protein could inhibit complement deposition of the streptococcal surface, and contribute to antiphagocytosis of M protein (Perez-Caballero et al., 2000). It has been suggested that the M proteins contain conserved fibrinogen binding sites in the B-repeat region (Whitnack et al., 1984) and the binding of fibrinogen blocks the activation of complement via the alternative pathway and reduces the amount of C3b bound to GAS, therefore reducing phagocytosis by PMNs (Whitnack and Beachey, 1982). It also has been suggested that M proteins of

streptococci inhibit chemotaxis of PMN and therefore can weaken defensive mechanisms of infected organism (Tylewska et al., 1978). In addition, Fischetti suggested that M proteins are structurally similar to α -tropomyosin, which is the prototype molecule for an α -helical coiled-coil fibrillar protein, and the similarity of the structure might contribute to anti-phagocytosis by interfering with the contractile proteins (Mische et al., 1987, Fischetti, 1989, Hosein et al., 1979).

The adherence of GAS to epithelial surfaces allows them to colonise host surfaces. Since Ellen and Gibbons found that M-positive GAS adheres well to human cheek epithelial cells *in vitro* whereas M-negative GAS do not (Ellen and Gibbons, 1972), streptococcal M protein has been implicated as an adhesin during GAS infection. Several studies have proved that M proteins are responsible for GAS binding to pharyngeal epithelial cells. Courtney and colleagues found that the isogenic M-positive strain could adhere to Hep-2 cells whereas a M-negative strain could not (Courtney and Hasty, 1991). Others found that the M-positive organisms adhered to pharyngeal epithelial cells more avidly than the M-negative organisms (Tylewska et al., 1981, Ellen and Gibbons, 1972, Tylewska and Hryniewicz, 1987). Although later it was found that M protein could not bind to buccal epithelial cells (Beachey and Ofek, 1976), there are several studies that prove that M proteins are important for GAS attachment to the host. For example, Okada and colleagues showed that CD46 on keratinocytes is the receptor for M proteins (Okada et al., 1995, Okada et al., 1994) and Perez-Casal et al. showed that the C-repeat region of M proteins is the domain that binds to keratinocytes (Perez-Casal et al., 1995). The mechanisms of adhesion to the pharyngeal and skin may be different (Cunningham, 2000). It has been suggested that the adhesion of GAS is complex, and the role of M proteins in adherence varies with the M protein serotype, the type of host cell and other environmental factors (Schrager et al., 1998).

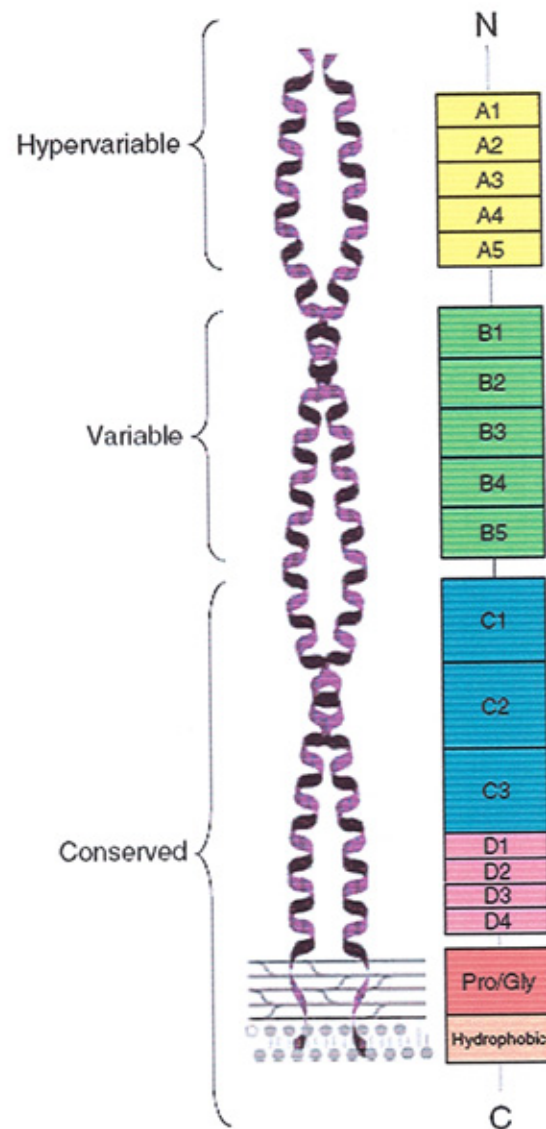


Figure 1.5 Structure of M proteins.

Blocks A, B, C, and D designate the location of the sequence repeat blocks. Numbers above the block indicate the number of repeated amino acids peptides per block. The C-terminal end is located within the cell wall and membrane (Bisno et al., 2003).

Additionally, a group of streptococcal surface proteins which are structurally similar to M proteins have been characterised and are termed M-like proteins. Functional studies have shown that M-like proteins play roles in GAS pathogenesis through their ability to interact with a variety of human proteins including plasminogen, Immunoglobulin G and Immunoglobulin A (Frithz et al., 1989, Gomi et al., 1990) and fibrinogen (Johansson et al., 2004). M-like proteins differ in the types of repeats and in their ability to interact with human protein (Cunningham, 2000).

1.2.3.2 The HA capsule

The HA capsule plays multiple functions in GAS, and it is an extracellular virulence factor. Figure 1.6 shows that GAS strains grow as large, spreading, wet colonies on solid media. This characteristic mucoid colony morphology is due to the abundant production of the HA capsule. The synthesis of HA is controlled by an operon consisting of three different genes called *hasA*, *hasB* and *hasC*, encoding HA synthase, UDP-glucose dehydrogenase and UDP-pyrophosphorylase, respectively (Crater and van de Rijn, 1995). Streptococcal strains vary in the amount of HA capsule they produce, although the *has* operon is highly conserved (Alberti et al., 1998).

The HA capsule produced by GAS may be used to alter their microenvironment to enhance survival of desiccation (Roberson and Firestone, 1992). Interestingly, the HA capsule is chemically similar to the HA that appears in the host connective tissue (Figure 1.2). Roberts suggested that polysaccharide capsules in general may mediate a number of biological processes, including invasive infections of human beings (Roberts, 1996). Cleary and Larkin reported that aggregation of GAS cells by their HA capsule provides a novel means to avoid self-destruction by oxygen metabolites (Cleary and Larkin, 1979). Several studies have shown that acapsular mutant strains had reduced their virulence and colonization capacities in animal models (Schmidt et al., 1996, Husmann et al., 1997, Wessels and Bronze, 1994, Wessels et al., 1991).

It has been reported that GAS strains rich in M protein and capsule are extremely virulent to humans, and mucoid strains have long been seen to cause deeply invasive infections (Bisno et al., 2003).



Figure 1.6 Muroid colonies of *Streptococcus pyogenes* on blood agar.
S. pyogenes causes β -hemolysis.

Although HA and M proteins are variously important in resistance of different GAS to phagocytosis. Foley and Wood showed that hyaluronidase treatment of encapsulated streptococci increases their susceptibility to phagocytosis, proving that the HA capsule is antiphagocytic (Foley and Wood, 1959). This was also supported by several other studies (Wessels et al., 1994, Moses et al., 1997, Dale et al., 1996, Whitnack et al., 1981). For example, Moses et al showed that expression of increased amount of capsule in a poorly encapsulated GAS strain enhance resistance to phagocytosis (Moses et al., 1997). However, Dale et al found that type 24 M protein which binds to fibrinogen plays a crucial role in resisting opsonization, whereas type 18 strain only can resist opsonization when HA was produced (Dale et al., 1996). The mechanism by which the HA capsule of GAS provides resistance to phagocytosis is not yet established. However, Dale et al. indicated that the ability of the GAS HA capsule to prevent phagocytosis is not due to the effect on C3 deposition on the surface of GAS, although the various complement deposition on different capsular types of *S. pneumoniae* have been shown to affect phagocytosis (Neeleman et al., 1999, Hostetter, 1986), because they demonstrate that encapsulated and unencapsulated M18 GAS were equally opsonized by C3 in either plasma or serum, yet only encapsulated organisms resisted phagocytic killing in blood (Dale et al., 1996). Cunningham suggested that reduced opsonisation is due to the physical barrier of the HA capsule preventing access of phagocytes to opsonic complement proteins bound the bacterial surface (Cunningham, 2000).

The HA capsule also plays a complex role in adherence to the epithelial cells, since it is able to bind to CD44 on epithelia cells even in poorly encapsulated strains. (Okamoto et al., 2004, Cywes et al., 2000, Schrager et al., 1998). Husmann et al. showed the capsule-deficient strains were impaired in their ability to colonized the pharynx in mice model and they also found that capsule was necessary for survival of GAS in the upper and lower respiratory tracts (Husmann et al., 1997). These evidences suggested that HA mediates attachment of GAS to pharyngeal and epidemial keratinocytes. However, as HA forms an outside layer on GAS cells, capsules may also prevent interaction of bacterial cell wall protein with host cell receptors. Kawabata et al. also found that an encapsulated wild type strain, SSI-1, adhered to and invaded into pharyngeal epithelial cells at lower frequencies when compared with their unencapsulated isogenic mutants, although in mouse lethality experiments, the

wild type strain exhibited 10^5 -fold higher virulence than the mutant (Kawabata et al., 1999).

In a recent study, Lida et al. compared the effect of wild type strain and *hasA*-knockout strains on delayed death of mice in a model of STSS, which is similar to human STSS, and they showed that when wild type strain was injected, 5 of the 10 mice died but all of the mice infected with mutant strain survived, so they conclude that the HA capsule was a critical pathogenic factor for the delayed death in the mice model of STSS (Iida et al., 2006).

Due to the HA capsule being structurally similar to that found in human connective tissue, it has been defined as weakly immunogenic and antibodies to GAS HA have been quite difficult to demonstrate in humans (Bisno et al., 2003).

1.2.3.3 Lipoteichoic acid

As a Gram-positive bacterium, the cell wall of GAS contains lipoteichoic acid. Since Beachey and Ofek found that GAS lacking M proteins adhered to human oral mucosal cells equally as well as GAS rich in M proteins. They suggested that lipoteichoic acid, rather than M proteins binds to epithelial cells via ester linked fatty acids (Beachey and Ofek, 1976). This was proved by Simpson and Beachey, as they found that fibronectin of oral epithelial cells is the receptor for GAS binding and this activity is inhibited by adding lipoteichoic acid (Simpson and Beachey, 1983). In Cunningham's review, she indicates that LTA accounted for approximately 60% of adhesion to epithelial cells (Cunningham, 2000). Also Courtney et al. found that human PMNs possess specific binding sites for lipoteichoic acid (Courtney et al., 1981), and later suggested that at least two streptococcal adhesins, lipoteichoic acid and M protein, are involved in the adherence of streptococci to certain cells and that the relative contributions of these adhesins to the attachment process depends on the type of host cells used to in the study (Courtney et al., 1992, Courtney et al., 1994b, Courtney et al., 1994a). Therefore Cunningham suggested in her review that environmental factors expressed in a particular body site may result in particular adhesins expressed by GAS (Cunningham, 2000), and Schrager et al. postulated that LTA mediates a weak initial attachment

followed by a stronger and more specific one such as M proteins or fibronectin binding proteins (Schrager et al., 1998).

1.2.3.4 Pyrogenic exotoxins

Most strains of GAS produce one or more toxin that are called streptococcal pyrogenic exotoxins (Spe), due to the fact that they are able to induce fever and reflect their identity as superantigens (Caparon, 2001). The Spes consist of several distinct antigenic types; among of them, SpeA, SpeB, and SpeC are well-defined. SpeA and SpeC are also called erythrogenic toxins as they are associated with the red rash observed in patients during scarlet fever (Cunningham and Watson, 1978). They are often encoded by the genes in the genomes of chromosomally integrated lysogenic bacteriophage (Johnson and Schlievert, 1984, Goshorn and Schlievert, 1989). SpeA and SpeC are genetically related to the *Staphylococcus aureus* enterotoxins, sharing significant primary sequence homologies with these staphylococcal toxins (Betley et al., 1992). They contain distinct T-cell receptor (V β region) (TCR) binding site and class II major histocompatibility complex (MHC) binding sites. They are able to simultaneously bind to the MHC class II and the TCR, which causes the T cells to produce a large amount of inflammatory cytokines (Bisno, 2001, Cunningham, 2000). Due to superantigenic activities SpeA and SpeC play a role in STSS.

Unlike SpeA and SpeC, the gene encoding for SpeB is chromosomally located, and all strains of GAS produce SpeB. The expression of SpeB is activated by the Rgg family member RopB (Chaussee et al., 1999) and a Rgg-like regulator RopA is involved in the secretion and folding of zymogen SpeB (Lyon et al., 1998). Kapur et al. found that genes encoding for SpeB and the streptococcal cysteine protease were identical (Kapur et al., 1993a, Kapur et al., 1993b, Gerlach et al., 1983). Streptococcal cysteine proteases are enzymes of proteolytic activity and cleave many host proteins to cause disease. SpeB has been reported to contribute to resistance to phagocytosis. It has been shown that the extracellular cysteine proteases produced by GAS are able to cleave human interleukin-1 β precursor to form biologically active interleukin-1 β , a major cytokine mediating inflammation and shock (Kapur et al., 1993b, Kapur et al., 1993a). Matsuka et al. reported that streptococcal cysteine protease were able to degrade

fibrinogen (Matsuka et al., 1999). Eriksson and Norgren showed that SpeB was able to cleave antigen-bound immunoglobulin G, and implied that this capability contributes to streptococcal persistence in opsonizing blood (Eriksson and Norgren, 2003). Another study showed that SpeB induced PMN cells to undergo apoptosis, subsequently reducing phagocytic activity in U937 cells. Chiang-Ni et al. in a recent study reported that SpeB could cause mitochondria damage to PMN cells, preventing immune clearance at an early infectious stage (Chiang-Ni et al., 2006). Extracellular cysteine proteases also have been shown to have the ability to degrade several human extracellular matrix proteins, including fibronectin and vitronectin (Kapur et al., 1993b). Destruction of extracellular matrix and cells would make it easier for the bacteria to spread through the host tissues. In addition to the ability to resist phagocytosis, SpeB is also responsible for the streptadhesin activity of GAS, i.e. the ability to bind glycoproteins such as thyroglobulin, fetuin, asialofetuin and submaxillary mucin (Hytonen et al., 2001). Also, SpeB has been found to be involved in STSS and/or streptococcal infection of skin and mucous membrane (Ohkuni et al., 2004). Ohkuni also showed that recombinant SpeB increased capillary permeability of pig skin through induction of the release of histamine from human mast cells. In addition, several reports have suggested that SpeB is able to change the GAS surface by cleaving M protein (Raeder et al., 1998), fibronectin binding protein (Matsuka et al., 1999, Nyberg et al., 2004) and the streptococcal C5a peptidase (Berge and Bjorck, 1995). For this reason, the contribution of SpeB to GAS pathogenesis is still controversial (Holm et al., 1992, Cunningham, 2000, Kansal et al., 2000, Gubba et al., 1998, Cu et al., 1998, Talkington et al., 1993).

1.2.3.5 Fibronectin-binding proteins

Fibronectin-binding proteins including M proteins, M-like proteins and protein F are cell wall anchored proteins, and they play an important role in GAS attachment and invasion. They contain an *N*-terminal signal sequence involved in transmembrane transport of the proteins, a *C*-terminal cell wall- and membrane-spanning region, which includes cell wall anchoring motif LPX[T,S,A]G and a short positively charged *C*-terminal intracellular tail (Schwarz-Linek et al., 2006). It has been shown that most GAS strain produce fibronectin-binding protein, and they all show high fibronectin-

binding ability. The Fibronectin-binding protein of GAS is directly involved in the fibronectin-mediated adherence of GAS to epithelial cells in the internalization process (Hynes, 2004). Natanson et al. suggested that the presence of protein F and the capacity to bind fibronectin correlated strongly with the M type of various strains of GAS. This correlation may suggest the existence of a relationship between fibronectin binding and the pathogenic potential of GAS (Natanson et al., 1995, Talay et al., 1992, Hanski and Caparon, 1992). Nakata et al. compared wild type strain and a MsmR (a positive regulator of the major fibronectin-binding adhesin protein F2 in a serotype M49 strain) negative mutant strain, and found that altered levels of fibronectin-binding proteins in the mutant decreased adherence to and internalization into human pharyngeal epithelial cells (Nakata et al., 2005). The role of fibronectin-binding proteins in streptococcal infection has been addressed in several studies via the deletion of the genes encoding fibronectin-binding proteins and results in dramatic decrease in mouse mortality (Terao et al., 2002, Courtney et al., 1999).

1.2.3.6 C5a peptidase

The C5a peptidase is another streptococcal surface protein, and presents on all strains of GAS. C5a peptidase is a proteolytic enzyme and specifically cleaves complement C5a at His-Lys, which stimulate phagocytes to leave the bloodstream and guides them to the infection site (Cleary et al., 1992). As a result, if the concentration of C5a is reduced, chemotactic signals for phagocytic cells decrease and phagocytic cells can not move to the site of infection. Ji and colleagues compared the virulence of wild type and mutants in the *scpA* gene, which encodes C5a peptidase, in a mouse air sac model of inflammation. They found that *scpA*-negative mutants were cleared more efficiently than wild-type bacteria and the detection of the location of bacteria revealed that mutants were transported to lymph nodes, whereas wild-type streptococci avoided transport to the lymph nodes and rapidly spread to the spleen (Ji et al., 1996). Also Husman and colleagues demonstrated that deletion of the *scpA* gene resulted in a small but significant difference from the wild type in the incidence of pneumonia in a mouse model (Husmann et al., 1997). All these studies proved that C5a peptidase contribute to GAS resistance of phagocytosis.

In addition to its proteolytic activity, C5a peptidase has been found to bind fibronectin (Beckmann et al., 2002). Beckmann et al. detected fibronectin binding activity of recombinant C5a peptidase and they also found that C5a peptidase-negative Group B streptococcus mutants show 50 % less binding activity to fibronectin than wild type, so they conclude that this protein has a second important role in streptococcal infection.

In a recent study, C5a peptidase was found to play a role in GAS invasion (Purushothaman et al., 2004). It is well known that M proteins contribute to invasion of epithelial cells by GAS, and this ability is independent of other fibronectin proteins, such as protein F. Since M1 protein is the major invasion of this serotype GAS, and contribute to 90 % of invasion, Purushothaman et al. generated M-negative mutants, and determined the role of C5a peptidase in GAS invasion of epithelial cells. They found that C5a peptidase produced by M-negative mutant promotes the invasion of organisms into Hep-2 cells, and this activity was at the same level in the absence or presence of fibronectin. They therefore concluded that C5a peptidase promoted fibronectin-independent GAS invasion.

1.2.3.7 Hyaluronate lyase

Most of the hyaluronate lyases, also incorrectly referred to as hyaluronidases (Smith et al., 2005), of GAS are encoded by bacteriophage genomes that have integrated into the host chromosome (Hynes et al., 1995). The ability of hyaluronate lyases to cleave HA, via β -elimination across the β -1-4-linkage of HA, and producing unsaturated disaccharides, results in a decrease in the viscosity of the extracellular matrix, which may increase permeability of the connective tissues and therefore potentially increase the pathogen aggression and diffusion of its toxins through the connective tissue. Hyaluronate lyases are consequently often referred to as the spreading factor. However, there is no experimental data to support this assumption. Several recent studies, for instance, have suggested that the role of hyaluronate lyases is to digest the HA capsule of the host organism (Hynes and Walton, 2000). Study by Smith et al. have suggested that the role of hyaluronate lyase produced by the M1 serotype GAS was unlikely to be for the complete degradation of connective tissue as the enzyme is only active against HA and not active against other glycosaminoglycan (Smith et al,

2005). Instead, the likely role of this enzyme, as it was shown to introduce widely-spaced cuts in HA, is to facilitate a local reduction in HA capsule viscosity to aid phage invasion of the streptococcal host. As this seems to be a conflict of interest with regards to the anti-phagocytosis of HA produced by GAS, Hynes (2004) argued that during hyaluronate lyase production, other antiphagocytic factors such as M protein provide sufficient protection for GAS. In addition, the latest study by Starr and Engleberg suggested that the function of hyaluronate lyase is in the specific breakdown of HA in tissue or even its capsule for nutritional purposes (Starr and Engleberg, 2006). Whether the hyaluronate lyase play a role in GAS infection is unknown, however the antibodies to the protein are detected following streptococcal infections (Halperin et al., 1987).

1.2.3.8 Streptokinase

Streptokinase, also known as fibrinolysin, is expressed by all strains of GAS. It is a plasminogen-binding protein on the bacteria surface (Reed et al., 1995). Once the host plasminogen is bound to the bacteria surface, forming a streptokinase-plasminogen complex, it is converted to plasmin by streptokinase. Plasmin is an active proteolytic enzyme that can dissolve blood clots and degrade extracellular matrix proteins, such as fibrin which limits the area of bacterial infection (Lottenberg et al., 1994). Therefore streptokinase is referred to as a spreading factor due to the ability to inhibit the deposition of fibrin barriers and facilitate movement of GAS through normal tissue barriers (Huang et al., 1989). It has been reported that the human specificity of GAS might be due to the fact that streptokinase has only limited ability to interact with plasminogen from species other than humans (Ringdahl et al., 1998). To study the mechanisms of streptokinase in animal model, Li et al. set up a two-stage strategy (Li et al., 1999). In the first *in vitro* stage, bacteria were preincubated with human plasma or plasminogen-depleted plasma in the presence or absence of exogenous streptokinase to enable assembly of a surface enzymatic complex. The treated bacteria were then washed free of unbound human proteins and tested for effects on virulence in a second *in vivo* stage (mouse model). They found that bacteria preincubated with human plasma could assemble a surface plasminogen-streptokinase complex that could activate mouse plasminogen, and by monitoring the time to death for infected mice, they concluded

that acquisition of a surface-associated plasmin(ogen)-dependent enzymatic activity can contribute to the virulence of GAS invasive infections. Rezcallah et al. analysed the streptokinase expression using the same strategy, injecting pretreated wild type M1 strain and more virulent mutant strain into mice skin, and found that the more virulent strain produced more streptokinase (Rezcallah et al., 2004). In 2005, Tadayoshi and colleagues compared streptokinase gene expression in *emm49*-genotyped GAS strains isolated from patients with severe invasive and non-invasive infections, and found that the expression of this gene in isolates of severe invasive infections was significantly higher than that in isolates of non-invasive infections, and they speculated that the production of streptokinase may be a causative factor for streptococcal toxic shock syndrome (Ikebe et al., 2005).

1.2.3.9 Streptolysins

Most GAS express two different haemolysins called Streptolysin O (SLO) and Streptolysin S (SLS). Both of these lyse erythrocytes, polymorphonuclear leucocytes and platelets by forming pores in their cell membrane (Greenwood et al., 1992). SLO is a member of the thiol-activated family of toxins (Fontaine et al., 2003), has oxygen lability (Alouf, 1980) and is irreversibly inhibited by cholesterol (Bisno et al., 2003). SLO damages cholesterol-containing membranes by binding to cholesterol, oligomerizing and inserting into the plasma membrane, resulting in the creation of large pores in the eukaryotic cell membranes. SLO is produced by almost all GAS strains and is antigenic. It plays a role in the pathogenesis of post-streptococcal rheumatic fever. In addition to its cytolytic activity, SLO has been shown that it mediates the translocation of GAS NAD-glycohydrolase (NADase) into human epithelial cells in vitro (Madden et al., 2001). These two proteins interact functionally as a compound signalling toxin. Limbago et al. have demonstrated that the mice injected with a SLO negative strain were more likely to survive than those infected with SLO positive strain (Limbago et al., 2000). Bricker et al showed that internalization of an encapsulated SLO negative mutant was 4-fold higher than the wild-type parent strain, which suggested that SLO and/or the exotoxin NADase contribute to GAS survival in the human host by inhibiting internalization of the bacteria into pharyngeal epithelial cells (Bricker et al., 2002). In a later study,

Hakansson et al. examined the role of the two exotoxins, SLO and NADase, in the process of internalization of GAS in primary keratinocytes and found that (i) SLO rather than NADase was primarily responsible for the inhibition of internalization and they also found that (ii) SLO blocks trafficking of internalized bacteria to lysosomes (Hakansson et al., 2005). In Brickers et al. later study, they used a mouse model of invasive soft-tissue infection to compare the relative virulence of an invasive clinical isolate of an M type 3 strain of GAS with that of isogenic mutants deficient in NADase or both NADase and SLO. They found that NADase enhances SLO-mediated cytotoxicity by inducing host-cell apoptosis and, by blocking the internalization and killing of GAS bound to the cell surface and SLO, in the absence of NADase, actually may mitigate GAS virulence (Bricker et al., 2005).

SLS is a non-antigenic haemolysin and it is responsible for the β -haemolysis around colonies on blood agar plates. It exists in intracellular, cell-surface bound and extracellular forms and is one of the most potent cytotoxins known (Bisno et al., 2003). Like SLO, SLS can damage the membranes of polymorphonuclear leucocytes, platelets, and subcellular organelles with its cytolytic ability. The experiments using animal model showed that SLS has lethal effects. Betschel et al. injected a SLS deficient mutant and a wild type GAS into mouse and found that SLS deficient mutants are markedly less virulent than their isogenic parents (Betschel et al., 1998). Miyoshi-Akiyama et al. injected a wild type GAS strain and a SLS-deficient mutant into mice and found that the SLS-deficient mutant had a 10-fold lower virulence level than did the parent strain (Miyoshi-Akiyama et al., 2005). They also analyzed the *in vivo* dynamics of peritoneal exudate cells (PECs) in mice injected with GAS strains (virulent and avirulent). The injection of a live avirulent strain or heat-killed organisms of the virulent strain significantly increased the number of PECs (primarily neutrophils), whereas a live high-virulence strain did not. Furthermore, coinjection of the virulent strain was shown to could suppress the neutrophil infiltration induced by thioglycolate and they suspected that this suppressive effect was mediated by cell-associated SLS that actively destroyed neutrophils recruited to the site of infection (Miyoshi-Akiyama et al., 2005). In addition, a cell-bound type of SLS has been shown to have strong cytotoxic effects on various host cells, including neutrophils (Ginsburg, 1999). These evidences showed that SLS damages host cells and contributes to the pathogenicity of GAS infection.

1.2.3.10 DNAase

GAS produces at least four distinct forms of DNAases, including DNAase A, B, C and D, among which DNAase B is the most common form (Bisno et al., 2003). DNase B is also designated as mitogenic factor and is synthesized by pathogenic streptococci during infection (Sriskandan et al., 2000). It has been suggested that all strains of GAS have extracellular nuclease activity and produce significantly more acidity compared to other groups of streptococci, which may contribute to virulence (Wannamaker and Yasmineh, 1967). Although it had been found that the level of anti-DNase B antibody in invasive cases was significantly lower than those in healthy individuals, which suggested that DNase B plays a role in invasive disease, the potent mechanisms of DNase has only been revealed in recent studies. Caparon suggested that this group of enzymes assists GAS in spreading through tissue by their abilities to hydrolyse nucleic acids that become viscous when released by damaged cells (Caparon, 2001). Chausse et al. discussed that host mucus may contain significant amounts of DNA that can inhibit the adherence of GAS to epithelial cells (Chaussee et al., 2001). DNAase produced by GAS may cleave this DNA to help the adherence of GAS. In addition they also suggested that secreted DNAase also cleave the DNA in pus to decrease the viscosity of pus and facilitate bacterial dissemination. Broudy et al. found that streptococcal coculture with human pharyngeal cells resulted in the induction of lysogenic bacteriophage as well as the phage-encoded DNase, and speculate that the induction of phage-encoded DNase enhancing the fitness of both bacteria and phage (Broudy et al., 2002). In 2005, Sumby and colleagues using *in vitro* human PMNs experiments demonstrated that one function of secreted GAS DNase is to protect against extracellular killing by PMNs, presumably through the degradation of neutrophil extracellular traps (NETs), and they used an animal infection model to confirm that extracellular DNase contributes to normal GAS disease progression (Sumby et al., 2005). Buchanan et al. demonstrated that the inhibition of GAS DNase activity with G-actin resulted in the clearance of more GAS by neutrophils *in vitro*, and reduced GAS virulence in a murine model of the invasive infection, necrotizing fasciitis (Buchanan et al., 2006). All the evidences showed that DNase play a role in invasive disease.

1.3 Studies of virulence factors expression in GAS

As mentioned previously, GAS produces a wide range of virulence factors. These factors play different roles in streptococci pathogenesis, and some of them play multiple roles. Probably for the reason to conserve energy, they may not be expressed at the same time, but they are only induced when required for infection, depending on the site of the infection and stage of the infection (Graham et al., 2001). Despite a large number of studies that have involved the analysis of the expression of the streptococci virulence factors, the molecular basis of many streptococci infections remains largely unknown. Classical approaches, such as the generation and characterization of isogenic mutants can only analyse small amount of bacterial gene products that are potentially involved in virulence at the same time. Therefore, post-genomics approaches such as transcriptomics and proteomics are being used to investigate the expression of virulence factors

1.3.1 Development of Genomics

The last decade in science belongs to genome sequencing. Since the first complete genome of a free-living organism, *Haemophilus influenzae* was published in 1995 (Fleischmann et al., 1995), the number of genome sequences has grown exponentially (Liolios et al., 2006) (Figure 1.7 and 1.8). Besides the human genome sequencing project (Venter et al., 2001), many other scientifically, medically and economically interesting organisms have been sequenced. As of September 2006, there were 429 complete genome sequences has been published, which includes 28 archaeal, 359 bacterial and 42 eukaryal. In addition, there are 2172 genome sequence projects currently ongoing, including 57 archaeal, 993 bacterial and 634 eukaryal (<http://www.genomesonline.org/>) (Figure 1.8).

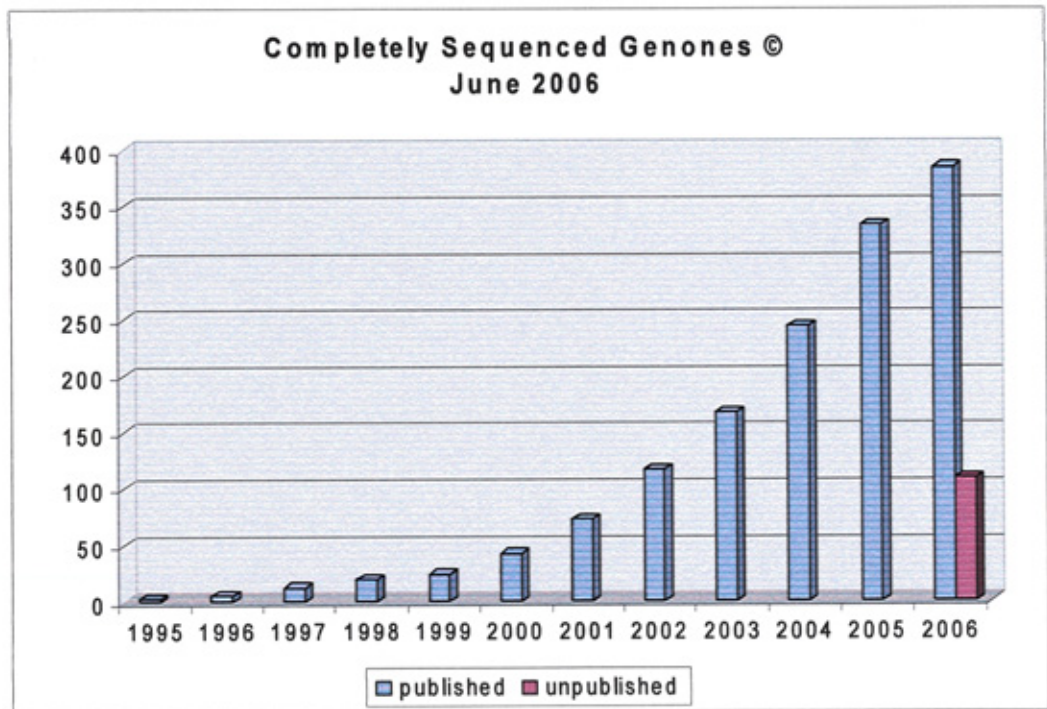


Figure 1.7 Completely Sequenced Genomes according to Genomesonline.

The published sequenced genome in the first half year of 2006 is more than total published sequence genome of 2005.

http://www.genomesonline.org/Gold_statistics.html

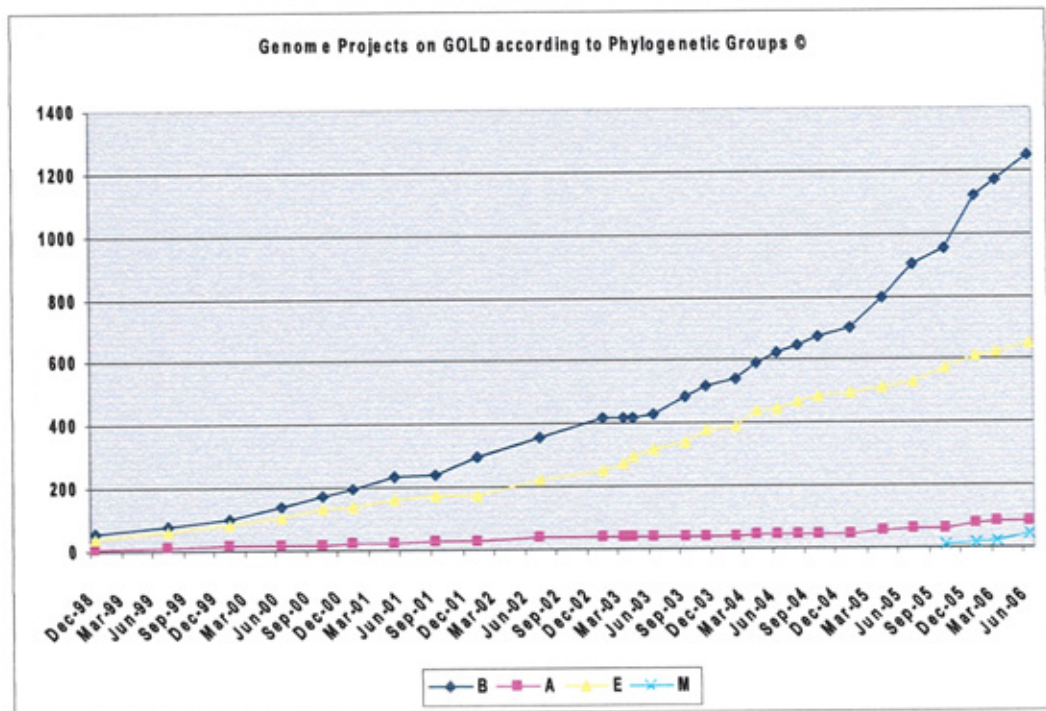


Figure 1.8 Genome Sequencing Projects According to Genomesonline.org.

Until September 2006, there are 2172 genome sequence projects currently undergoing, including 57 archaeal, 993 bacterial and 634 eukaryal.

http://www.genomesonline.org/Gold_statistics.html

As genome sequence information has accumulated, the paradigm has shifted from sequencing to identification and functional analysis. The sequence information is stored so as to facilitate access for the study of biological function. Generally, after the genome of a new organism is sequenced, the raw DNA sequence will be annotated as putative open reading frame (ORF), which contains protein-coding regions. The amino acid sequences derived from ORFs are then searched against sequence database to determine the relationship to previously sequenced genes and their functions. When the search results in 'hits' to a gene encoding a known protein, the newly sequenced gene is generally annotated as a homolog of the best hit (Schwartz, 2000). This homology searching and ORF identification have been successfully used in comparative studies of closely related bacterial species and different strains of the same species (e.g. virulence and avirulence). However, around 40 % of identified ORFs can not be identified by this approach (Figure 1.9). In addition, due to gene rearrangements, one gene can code for more than one protein. Thus, knowing the sequence of the genome does not mean that the functions of proteins encoded by the genome are immediately recognised (Pennington and Dunn, 2001).

1.3.2 Large-scale analysis of streptococcal virulence factors

In the last few years, several large-scale analysis strategies have been developed to identify essential and pathogenicity-related genes in pathogenic microorganisms. One strategy is to identify genes via transposon-mutagenesis (Hamer et al., 2001). Basically, a pool of gene mutations is generated by random insertions of transposons into the genome of an organism. The resulting mutants are screened and the genes are identified by PCR. Signature-tagged mutagenesis (STM) is a valuable approach to identify genes involved in bacterial virulence (Hamer et al., 2001). STM uses a series of transposons that differ in that each transposon has a unique "PCR-amplifiable" sequence. Random insertion mutants are constructed, with the series of transposons, arrayed, then pooled and inoculated into the host organism. The "PCR-amplifiable" sequence of survival isolates are amplified and then hybridized to the array of original mutants. The lack of certain sequences in the array correlate to the mutants that could not survived in the host, and therefore the assumption is that the disrupted gene is

critical for survival in the host. The regions flanking the insertions in mutants of interest are then sequenced and compared to the genomic sequence for identification.

In addition to signature-tagged mutagenesis, gene fusion technologies have been developed. *In vivo* expression technology (IVET) is a strategy for identifying virulence-related genes active *in vivo* but relatively inactive *in vitro* (Mahan et al., 1993). The basic process of IVET includes (i) the mutation of the strain of the pathogen of interest in a gene encoding an essential growth factor (EGF); (ii) the construction of a plasmid carrying a promoterless EGF gene and a linked reporter gene (*rep*); (iii) the cloning of random fragments of the pathogen of interest upstream of the promoterless *egl-rep* synthetic operon; (iv) the transfer of the *egl-rep* synthetic operon constructs into the *egf* mutant strain with chromosomal integration to form a library of merodiploids; (v) the injection of the library into the host; and (vi) the correlation of survival strains with the promoters of genes expressed during infection, and therefore crucial for infection. To eliminate strains containing constitutively active fusions, survival mutants are screened for expression of the linked reporter gene (*rep*) on general growth media. Several variations on the IVET have been developed. Auxotrophy complementation-based selections are achieved using fusions to a promoterless *purA* (a gene encoding purine auxotrophs) (Mahan et al., 1993); antibiotic selections are achieved using fusions to a promoterless antibiotic gene (Mahan et al., 1995); dual reporter selections which use reporter genes such as *hly* (pore-forming hemolysin listeriolysin O) which play functions in *in vivo* selection of active gene fusions and screening of these fusions at a later stage (Gahan and Hill, 2000); and recombinase-based IVET (RIVET) selection is achieved via fusions to a promoterless resolvase gene, such as *tnpR*, the product of which is specific for the target sequence, *resI*, and therefore slices out a previously genome-integrated antibiotic resistance gene, which is flanked by two *resI* sites, resulting in the loss of antibiotic-resistance (Merrell and Camilli, 2000, Angelichio and Camilli, 2002, Rediers et al., 2005).

Differential fluorescence induction (DFI) is another gene fusion technology. It utilizes the green fluorescent protein (GFP) as a reporter to monitor promoter activity induced during infection, by cloning a library of DNA fragments upstream of promoterless *gfp*. The mutants are then sorted using fluorescence-activated cell sorting (FACS) based on

their fluorescence level, and the sorted library plated for single colonies and further screening (Marra et al., 2002, Rediers et al., 2005).

Since these large-scale strategies have been developed, they were widely applied to streptococcal studies. STM were used to study the virulence factors of *S. pneumoniae* in a murine model of pneumonia and bacteraemia (Polissi et al., 1998, Lau et al., 2001, Hava and Camilli, 2002), *S. agalactiae* in the neonatal rat sepsis model (Jones et al., 2000) and *S. sanguinis* in a rabbit model of endocarditis (Paik et al., 2005). Smith et al. identified seven genes induced *in vivo* via IVET, which were potentially involved in the pathogenesis of *S. suis* infection of piglets (Smith et al., 2001). Marra et al. identified, using DFI technology, 78 unique sequences of *S. pneumoniae* that were induced under various *in vitro* conditions, which were used to mimic aspects of the *in vivo* environment once the pathogen enters a host (Marra et al., 2002).

Although these large-scale strategies have contributed to understanding streptococcal pathogenesis, they also have limitations; neither transposon-mutagenesis nor gene fusion technologies are able to detect repressed genes during infection, nor none of these technologies could identify the genes expressed constitutively with products activated during infection.

1.3.3 Transcriptomics

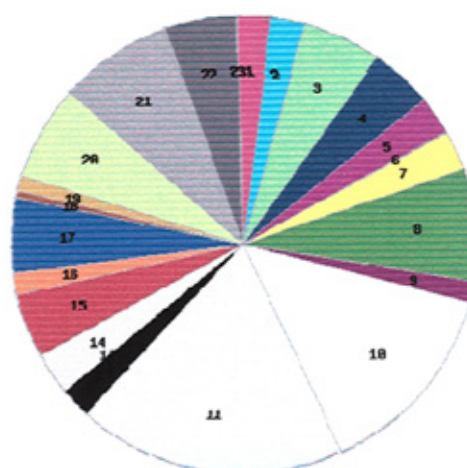
Transcriptomics has grown out of genome sequencing. This new approach which investigates biological function, is based on mRNA levels. The identification of which mRNA molecules are expressed in a cell in one condition infers which proteins are present at that condition. In transcriptomics, PCR products from cDNA (a complementary DNA copy of mRNA) or oligonucleotides are spotted onto a solid support. Target RNAs are extracted from the experimental samples labelled fluorescently and hybridised to genes. This results in the quantitation of each expressed mRNA or DNA from the sample that has a corresponding spot on the array (Kellam, 2001). The use of cDNA and oligonucleotides microarrays for high throughput analysis of mRNA expression on a genome-wide scale is having a dramatic impact on the investigation of gene function (Hughes et al., 2000). However, this

approach has some limitations. Firstly, there is poor correlation between mRNA and protein expression level (Gygi et al., 1999), and mRNA does not always reflect protein abundance. Additionally, the activity of the proteins encoded by mRNAs is regulated at several high levels beyond their mRNA or protein expression. It is known that, there are more than 400 possible chemical modifications of the products of mRNA. Such modification can be caused by addition of other molecules, for example by phosphorylation, glycosylation, protein processing and proteolytic cleavage (Pennington and Dunn, 2001). As a result, each transcript can give rise to more than one product and the measurement of mRNA expression cannot reveal all of these possibilities. This supports the hypothesis that even if all the genes of a genome have been sequenced, translated into proteins and analysed at the mRNA levels, it remains a major task to identify and characterize the various functional forms of each protein, in a particular cell type under a certain condition. Therefore detailed and direct analysis of the expression and structure of the proteins encoded by genomes are required to fully understand the biological process.

1.3.4 Proteomics

In 1995, a new fundamental concept called the proteome (PROTEin complement to a genome; http://ca.expasy.org/proteomics_def.html) emerged. The proteome comprises all the proteins that are expressed at a particular time by a particular cell, tissue or organism. Wsinger et al. defined the proteome as the 'total protein complement of a genome' (Wasinger et al., 1995). It is known that the genome contains static information whose content does not change with time. In contrast, the proteome is dynamics as the expression of proteins changes from moment to moment according to the development and physiological state of cells.

The new discipline, 'proteomics', which derives from the term proteome, was developed. Proteomics can be defined as 'the qualitative and quantitative comparison of proteomes under different conditions to understand cellular mechanisms underlying biological processes'(Anderson and Anderson, 1996).



Color	Gene Role	# of Genes	% out of 540419 Genes
1	Amino acid biosynthesis	11735	2.17%
2	Biosynthesis of cofactors, prosthetic groups, and carriers	12750	2.35%
3	Cell envelope	29720	5.49%
4	Cellular processes	23156	4.28%
5	Central intermediary metabolism	14218	2.63%
6	Disrupted reading frame	718	0.13%
7	DNA metabolism	15137	2.80%
8	Energy metabolism	42666	7.89%
9	Fatty acid and phospholipid metabolism	7927	1.46%
10	Hypothetical proteins - Conserved	75366	13.9%
11	Hypothetical Proteins	101305	18.7%
12	Other categories	11532	2.13%
13	Pathogen responses	2	0.00%
14	Protein fate	16320	3.01%
15	Protein synthesis	24260	4.48%
16	Purines, pyrimidines, nucleosides, and nucleotides	8165	1.51%
17	Regulatory functions	28830	5.33%
18	Signal transduction	2645	0.48%
19	Transcription	5835	1.07%
20	Transport and binding proteins	36420	6.73%
21	Unclassified	43485	8.04%
22	Unknown function	27124	5.01%
23	Viral functions	1103	0.20%

Figure 1.9 The percentage of genes in each role across all published genomes.

The pie chart shows the percentage of genes across all genomes in the Comprehensive Microbial Resource that are represented in each of the role families. The table under the pie chart shows the number and percentage of genes in each family represented. All role category data shown in this figure was generated on the TIGR (The Institute for Genomic Research) annotation of the genome and the data shown was generated at the time each of the genomes was entered into the CMR.

<http://210.218.199.237/tigr-scripts/CMR2/GenePieChart.spl?db=CMR>

The advent of proteomics has been dependent on the success of whole genome sequence projects, which ‘provide an essential platform for a wide range of complementary experimental approaches that ultimately support the understanding of how the products of these genes act together to regulate the activity of the organism’ (Pennington and Dunn, 2001). While genomics and transcriptomics provide basic information on DNA sequence and gene expression, proteomics provides quantitative and qualitative information on the total protein profile. In quantitative aspect, proteomics can evaluate the level of protein expression. In qualitative aspect, proteomics can determine the presence of protein isoforms and post-translational modifications. In addition, proteomics can examine the complexes of interacting proteins.

Since the development of techniques in protein separation using two dimensional gel electrophoresis (2DE) (O’Farrell, 1975) and electrospray ionization (ESI) for mass spectrometry (MS), which is a soft ionization method and can generate ions from large, non-volatile analytes, such as proteins and peptides, without significant analyte fragmentation (Kolker et al., 2006), 2DE coupling with ESI MS has become a major and popular method for proteomics studies.

1.3.5 Challenge in bioinformatics

In proteomics studies, bioinformatics play an important role in the identification and functional analysis of proteins. Bioinformatics began as the development of alignment tools including BLAST (Basic Local Alignment Search Tool) and FASTA, which are used to understanding the relationship between the sequences of genes or proteins (peptides). The programs can be freely accessed on NCBI website (<http://www.ncbi.nlm.nih.gov/BLAST/>). As genomics, transcriptomics and proteomics techniques developed, large amount of data related to the genes and proteins were generated. It is necessary to analyze these data in the context of the previously known information about location, structure and functions of genes or proteins to achieve the global understanding of biological systems. Therefore, bioinformatics has become a fundamental tool for processing these data (Oehmen et al., 2006). However, obtaining and interpreting biological knowledge from previously known information is an

arduous task. Consequently, development of automatic methods to assist in functional interpretation is one of the main changes in bioinformatics (Chagoyen et al., 2006)

As the M1 serotype of GAS, strain SF370 been completely sequenced in 2001 (Ferretti et al., 2001), many researchers have investigated the expressions of GAS virulence factors using transcriptomics (Smoot et al., 2001, Voyich et al., 2003, Graham et al., 2005, Virtaneva et al., 2005) and proteomics (Lei et al., 2000, Chaussee et al., 2001, Thongboonkerd et al., 2002, Chaussee et al., 2004, Aziz et al., 2004, Nakamura et al., 2004, Cole et al., 2005, Johansson et al., 2005). These studies help us to understand more about GAS infections, and help identify GAS infection drug target. Due to the powerful nature of these techniques more studies are ongoing.

1.4 Research Objectives

To investigate the pathogenesis of streptococci, I set up a simple biological system that was shown to reflect some elements of streptococcal infection. The system involved the propagation of streptococci in brain heart infusion media (control media) and brain heart infusion media with 0.5 % (w/v) HA (HA-rich media), as the presencet of HA, which is structurally identical to the HA in human connective tissue, may reflect streptococcal infection to a certain extent and therefore facilitate the identification of GAS proteins that have been upregulated. To develop proteomics technology on site and used this to examine simultaneous changes in multiple virulence factors in this simple system of streptococcal pathogenesis.

Specific objectives of this study are as follows:

- 1) To analyse the growth of the M1 serotype of GAS strain SF370 in control media and HA-rich media
- 2) To develop 2DE methods to separate the cell-associated proteins and extracellular proteins from streptococci.
- 3) To compare the protein expression of streptococci using 2DE.

- 4) To develop methods to analyse and identify protein expression using MS
- 5) To identify the differentially expressed proteins from M1 serotype of GAS, SF370 growing in control media and HA-rich media using 2DE in combination with the MS method developed in these studies
- 6) To functionally analyse differentially expressed proteins from the M1 serotype of GAS SF370 which may be involved in pathogenesis of GAS using a short signature search tool, TSSview (Turbinia/Biosystem Informatics Institute)

2 Material and Methods

The sources of all compounds is given in the Appendix A1

Unless otherwise stated, all solutions and buffers were made up with 18.2 Ω/cm H_2O purified by a Milli – Q Plus 18.2 H_2O purification system, and stored at room temperature.

Unless stated otherwise, all solutions and apparatus were sterilised by autoclaving (Appendix B1) at 121°C , with a pressure of 1.05 bar for 20 min. Solutions that could not be autoclaved were filter sterilised through a 0.2 μm Ministart[®] filter unit (Sartorius).

2.1 Media

Unless otherwise stated, all media were prepared using distilled H_2O , and sterilised by autoclaving. Liquid media were stored at room temperature, and solid media at 4°C . All adjustments to the pH of solutions are stated and performed using a pH meter (Appendix B2) and HCl or NaOH.

2.1.1 Luria- Bertani (LB) Broth

Per Litre

Tryptone	10.0 g
Yeast extract	5.0 g
NaCl	10.0 g

The pH of this medium was adjusted to 7.0 with NaOH, and then the medium was autoclaved before storing at room temperature.

2.1.2 LB agar

Per 100ml

LB broth	100 ml
Agar (bacteriological agar N° 1)	2.0 g

Agar was then made soluble by autoclaving the media and poured into Petri dishes when the media had cooled to less than ~50 °C.

2.1.3 Columbia Agar with 5 % (v/v) horse blood

Per 100ml

Columbia Agar powder	3.7 g
Horse Blood	5 ml

Agar was made soluble in 95 ml of distilled H₂O by autoclaving. When the media was cool down to ~50 °C, 5 ml of horse blood was added in and mixed gently. Then it was poured into Petri dishes. The fresh horse blood was stored at –20 °C in 5 ml aliquots and left at room temperature to thaw just before use.

2.1.4 Defined medium

The defined medium was made up according to the method used by Homer et al.. (Homer et al., 1993).

Basal medium

Per Litre

Sodium acetate	6 g
Casein hydrolysate	5 g

$\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$	7.94 g
$\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$	2.66 g
Ammonium sulphate	0.6 g
KH_2PO_4	0.44 g
K_2HPO_4	0.3 g
Trisodium citrate	0.225 g
Adenine	0.03 g
Uracil	0.03 g
Guanine	0.02 g
Ferrous sulphate	0.01 g
$\text{MnSO}_4 \cdot 4 \text{H}_2\text{O}$	0.0145 g
NaCl	0.01 g

One litre of basal medium was made up and divided into 100ml aliquots in 200 ml medical flasks. Then it was autoclaved, and stored at room temperature.

Vitamins mixture

Per 20 ml

Nicotinamide	20 mg
Pantothenate	8 mg
Pyridoxamine	8 mg
Riboflavin	4 mg
Thiamine	4 mg
Biotin	0.5 mg
Folic acid	0.1 mg
<i>p</i> -Aminobenzoate	0.1 mg

20 ml of vitamins mixture was made up and filter sterilised. Then the stock was stored at -20°C in 1.5 ml aliquots until use.

Other stocks

Per Litre

MgSO ₄	25 g
Cysteine- HCl	220 g
Na ₂ CO ₃	40 g
Glucose	18.01 g

The above four components were made up to 20 ml, and filter sterilised separately. Then the stocks were stored at room temperature.

2.1.4.1 Defined medium with glucose

Per 20 ml

Defined medium

Basal medium	18 ml
Vitamins mixture	20 µl
MgSO ₄	200 µl
Cysteine- HCl	200 µl
Na ₂ CO ₃	200 µl

Various volumes of glucose were aseptically added to the defined medium to give concentrations of glucose in the medium as shown in Table 2.1

All solutions were transferred using pipettes with sterilised tips. The media, Ga, Gb and Gc, were made up just prior to use.

Medium	Concentration of glucose (w/v)	Volume of glucose stock (µl)
Ga	0.02%	200
Gb	0.1%	1000
Gc	0.2%	2000

Table 2.1 Three defined media with various concentrations of glucose

2.1.4.2 Defined medium with sodium hyaluronate

Various volumes of sodium hyaluronate were added to the basal medium, prior to autolaving, to give concentrations HA in the medium as shown in Table 2.2, and they were autoclaved.

Ha, Hb and Hc medium were autoclaved before adding the following solutions:

Vitamins mixture	20 µl
MgSO ₄	200 µl
Cysteine- HCl	200 µl
Na ₂ CO ₃	200 µl

Sterile distilled H₂O was added to make each medium (Ha, Hb and Hc) up to 20 ml and all solutions were transferred using Gilson pipettes with sterilised tips. The Medium Ha, Hb and Hc were made up just prior to use.

2.1.5 Brain Heart Infusion Broth (BHI) and BHI with 0.5 % (w/v) sodium hyaluronate* Broth (BHI+HA)

Per L

Brain Heart Infusion powder	39 g
-----------------------------	------

The powder was dissolved in distilled H₂O and 80 ml aliquoted into 100 ml medical flask. To each aliquot, 0.4 g of sodium hyaluronate was added when appropriate. The aliquots were autoclaved prior to use or storage.

* Sodium hyaluronate is derived from bovine tracheal tissue, and molecule mass range is between 60-100 kDa.

Medium	Concentration of hyaluronan (w/v)	Sodium hyaluronate (mg)
Ha	0.02%	4
Hb	0.1%	20
Hc	0.2%	40

Table 2.2 Three defined media with various concentrations of hyaluronan

2.1.6 Cryogenic storage of bacterial stocks

Per 100 ml	
100 % (v/v) glycerol	50.0 ml

Bacterial stocks were created by mixing 0.5 ml of the culture broth with 0.5 ml of 50 % (v/v) glycerol. These stocks were then stored at -80°C .

2.2 Growth of bacterial

2.2.1 Bacterial Strains

E. coli TOP10 (Invitrogen)

Streptococcus pyogenes SF370 (ATCC 700294)

2.2.2 Storage of bacterial stocks

2.2.2.1 Storage of *E.coli* TOP10

E. coli TOP10 was stored in 50 % (v/v) glycerol at -80°C . (section 2.1.6)

2.2.2.2 Storage of streptococci

S. pyogenes SF370 was maintained at -20°C as Lenticules (Codd et al., 1998).

2.2.3 Growth of the bacteria on the agar plates

As mentioned in section 2.2.2, the bacterial stocks were maintained either at -20°C or -80°C for long term storage. Two days before setting up liquid cultures, the bacteria

were plated or streaked onto agar plates. Then bacteria were substreaked one day before setting up the liquid cultures.

2.2.3.1 Plating *E. coli* TOP10

Prior to plating, LB agar plates were surface dried by placing them open faced down at 65 ° C for 10 min. A glass spreader was sterilised by immersion in 70 % (v/v) ethanol, followed by removal of excess ethanol by passing quickly through a blue Bunsen flame, and used to spread 50 µl of bacterial suspension evenly over the surface of the agar. When all the liquid had been absorbed into the agar, the plate was incubated inverted at 37 ° C for overnight.

2.2.3.2 Streaking GAS

Prior to streaking, Columbia agar with 5 % (v/v) of horse blood plates were surface dried by placing open face down at 65 ° C for 10 min. One Lenticule was place on the surface of agar, and left for 5 min to melt. The melt bacteria were streaked on the agar plate using a sterile wire loop. Then the plate was incubated inverted at 37 ° C for overnight.

2.2.4 Growth of *E. coli* TOP10 for protein extraction

E. coli TOP10 was inoculated from an agar plate using a sterile wire loop, and grown in 50 ml of LB broth. The culture was grown overnight at 37 °C with orbital shaking at 100 rpm. The growth of bacteria was monitored by measuring the optical density (OD) at 600 nm, and was harvested at late exponential phase, an OD of 1, according to the sample preparation in 2DE protocols on website of The ExPASy (Expert Protein Analysis System) proteomics server (<http://www.expasy.org/>). The OD of the culture was checked by transferring 1 ml of culture into a cuvette and reading the OD in a spectrophotometer (Appendix B3) previously zeroed with LB media and set to 600 nm.

2.2.5 Growth of *S. pyogenes* SF370 in defined media

S. pyogenes SF370 was inoculated from a Columbia agar with 5 % (v/v) horse blood plate using a sterile wire loop, and grown in 20 ml of each of following media: Ga, Gb, Gc, Ha, Hb, and Hc (see section 2.1.4). The static cultures were placed at 37 °C, and the growth of bacteria was monitored every 24 h by measuring the OD at 660 nm. The OD of the culture was checked by transferring 1 ml of each culture into a cuvette and reading the OD in a spectrophotometer previously zeroed with corresponding medium.

2.2.6 Growth of streptococci in BHI and BHI + HA

S. pyogenes SF370 was inoculated from a Columbia agar with 5 % (v/v) horse blood plate using a sterile wire loop, and grown in BHI and BHI + HA. The static cultures of *S. pyogenes* were placed at 37 °C in a static incubator (Appendix B4).

For the purpose of monitoring the growth of GAS, the OD at 660 nm was monitored every hour by transferring 1 ml of each culture to a cuvette and reading the OD in a spectrophotometer previously zeroed with corresponding medium.

2.3 Investigation of the expression of hyaluronate lyases by *S. pyogenes* SF370

The analyses were based on the fact that if any hyaluronate lyases were produced by *S. pyogenes* SF370 when cultured, the degradation of hyaluronan would occur in the appropriate conditions if *S. pyogenes* culture supernatant or cell free extract was used as an 'enzyme' preparation. To determine if any hyaluronate lyases were expressed a 232 nm assay and a 3, 5-Dinitrosalicylic acid (DNSA) reducing sugar assay were employed.

2.3.1 Preparation of ‘enzyme’ for the assays

The cultures of *S. pyogenes* SF370 in BHI and BHI + HA (section 2.2.5.2) were centrifuged at 8,000 x g for 15 min (Appendix B5). The supernatants were saved as an ‘enzyme’ preparation to be used in the assays, and were labelled A and B, from BHI + HA and BHI respectively (Table 2.3).

The cells pellet was resuspended in 5 ml of 50 mM Tris-HCl (pH7.5), and the suspension was sonicated on ice at amplitude 14 microns for 1 min in 10 s bursts with 10 s intervals (Appendix B6). The lysate was then centrifuged at 14,000 x g for 30 min. The supernatant containing cell free extract was saved as an ‘enzyme’ preparation to be used in the assays, and were labelled C and D, from BHI + HA and BHI respectively (Table 2.3).

2.3.2 DNSA reducing sugar assay

The reducing sugar content of solutions was determined calorimetrically using DNSA (Miller, 1959)

2.3.2.1 DNSA reagents

Per Litre

3,5-Dinitrosalicylic Acid	10.0 g
Phenol	2.0 g
NaOH	10.0 g

Enzyme	Source	Culture
A	Supernatant	BHI+HA
B	Supernatant	BHI
C	Cell free extract	BHI+HA
D	Cell free extract	BHI

Table 2.3 Sources used in the detection of the expression of hyaluronate lyase

Per 10 ml	
Sodium sulphite	0.5 g

Per 10 ml	
Glucose	2.0 g

2.3.2.2 Reactions for DNSA assays

The reactions for DNSA assay were set up as follows:

Dialysed sodium hyaluronate * (2.66 mg/ml)	1.128 ml
Tris- HCl buffer (pH 6.5, 1M)	30 µl
Bovine serum albumin (BSA) (10mg/ml)	150 µl
18.2 MΩ/cm H ₂ O	177µl
Enzyme	15 µl

* The dialysed sodium hyaluronate was made following the protocol in Appendix C

The reactions were prepared for two time points. For the zero time point, ‘enzymes’ were added into the reaction, and mixed by vortexing. The reaction mixtures were boiled for 5 min immediately to denature the enzyme, and maintained at 37 °C overnight. For the overnight time point, enzymes were added into the reaction, and mixed by vortexing. The tubes were then incubated at 37 °C overnight. There were eight samples in total (Table 2.4).

To stop the reaction, 150 µl of each reaction was transferred into a clean microfuge tube and 150 µl of DNSA reagent was added. Then the tubes were boiled for 20 min, placed on ice for 10 min and finally equilibrated to room temperature before reading the absorbances. The absorbances were determined at 570 nm in a 96-well plastic microtitre plate (200 µl of each reaction was measured) using a plate reader (Appendix B7).

Sample No.	Enzyme	Reaction time
1	A	0
2	B	0
3	C	0
4	D	0
5	A	Overnight
6	B	Overnight
7	C	Overnight
8	D	Overnight

Table 2.4 The eight samples used in the DNSA assay

Reducing sugar release was determined from the standard curve (Figure D1). The standard curves were generated by the addition of 0-1.0 mg/ml of glucose (final concentration; 10 mg/ml stock concentration) of soluble sodium hyaluronate in 20 mM of Tris-HCl pH6.5, containing 1 mg/ml BSA (final concentration) to a final volume of 150 μ l. DNSA reagent (150 μ l) was added to the standard curve and processed in the same way as the enzyme assays (post-incubation).

2.3.3 232 nm assay

Changes in absorbance at 232 nm is consistent with the action of a lyase (β -eliminase), the action of which results in the production of a Δ 4,5 unsaturated product (Smith et al., 2005) (Figure 2.1).

The reactions for 232 nm assays were set up as follows:

Dialysed sodium hyaluronate (2.66 mg/ml)	187.5 μ l
Tris- HCl buffer (pH 6.5, 1M)	10 μ l
BSA (10mg/ml)	20 μ l
18.2 M Ω /cm H ₂ O	277.5 μ l
‘Enzyme’	5 μ l

All reactions were performed in a total volume of 500 μ l, and all components except the ‘enzyme’ were pre-warmed to 37 °C. Assays were performed spectrophotometrically at 37 °C at 232 nm by using a 1-cm pathlength quartz cuvette. The spectrophotometer was pre-zeroed against H₂O. The absorbances of reaction without enzymes A and B (Table 2.3) were collected for 2 min. Then 5 μ l of each enzyme was added respectively, mixed immediately, and data collected for 10 min.

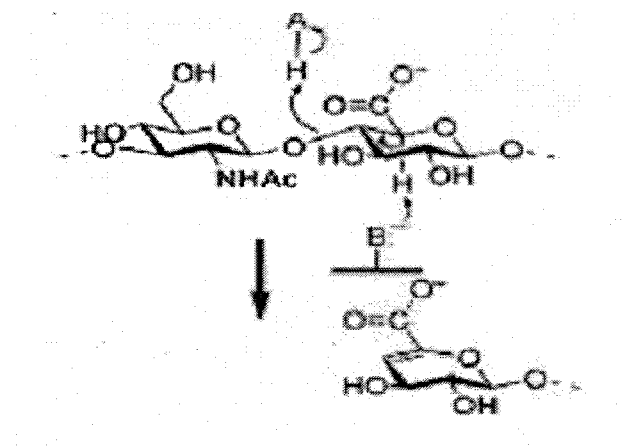


Figure 2.1 The β -elimination reaction catalyzed by a hyaluronate lyase (Smith et al., 2005).

This reaction results in the formation of a $\Delta 4, 5$ unsaturated product, which can be detected spectrophotometrically at 232 nm.

2.4 Investigation into the consumption of carbohydrates in *S. pyogenes* SF370 cultures

2.4.1 Preparation of samples for assays

The cultures of *S. pyogenes* SF370 in BHI and BHI + HA (section 2.2.5.2) were centrifuged at 8,000 x g for 15 min. The supernatants were saved as samples to be used in the assays. BHI and BHI + HA without inoculation were used as controls respectively. Four reactions were set up (Table 2.5).

2.4.2 DNSA reducing sugar assay

DNSA reducing sugar assay was employed to check the total reducing sugar consumption in *S. pyogenes* SF370 cultures. DNSA reagents were prepared as section 2.3.2.1, and the assays were performed as the protocol in section 2.3.2.2. The standard curve was generated by the addition of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0 mg/ml glucose (final concentration) to 18.2 MΩ/cm H₂O (Figure D2)

2.4.3 Analysis of the uronic acid consumption in the cultures

Since HA contains a uronic acid moiety, the determination of the change in uronic acid content in *S. pyogenes* SF370 cultures is an indirect method of determining the consumption of hyaluronic acid. A microtiter plate assay was employed for the determination of uronic acids (van den Hoogen et al., 1998)

Sample No.	Enzyme	Reaction time
1	A	0
2	B	0
3	C	0
4	D	0
5	A	Overnight
6	B	Overnight
7	C	Overnight
8	D	Overnight

Table 2.5 The four reactions that were set up for analysis of consumption of carbohydrates in *S. pyogenes* SF370 cultures

A 96-well plate was used for the assay. To each well 40 µl of sample (section 2.4.1) was applied and 200 µl concentrated sulfuric acid [96% (v/v)] containing 120 mM sodium tetraborate was carefully added. After mixing of the sample and the reagent by trituration with a pipette tip, the plate was placed in an incubator for 1 h at 80°C. Then 40 µl of metahydroxydiphenyl reagent [100 µl of metahydroxydiphenyl in dimethyl sulfoxide (100 mg/ml), mixed with 4.9 ml 80% (v/v) sulfuric acid just before use] was added and mixed with the samples as described above. After 15 min the absorbance of the pink-coloured samples was read at 570 nm, using an EL808 Ultra microplate reader (BIO-TEK Instruments Inc, Appendix B7). For each microtiter plate, 40 µl of the glucuronic acid standard series (0, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 µg) was treated as above at the same time (Figure D3).

2.5 2DE of *E. coli* TOP10

2.5.1 Solutions for protein extraction from Gram-negative bacteria

Phosphate Buffered Saline (PBS) (1 ×)

Per Litre

NaCl	8 g
KCl	0.2 g
KH ₂ PO ₄	0.2 g
Na ₂ HPO ₄ .12 H ₂ O	1.44 g

The pH of PBS stock was adjust to 7.4, and was autoclaved before storing at room temperature.

Lysis solution A stock

Per 10 ml

Urea	4.8 g
CHAPS	0.2 g

10 ml of lysis solution A stock was made in sterile universal and was aliquot into 100 μ l to store at -20°C . 2 μ l of IPG buffer was added just prior to use to make final concentration of 2 % (v/v).

Bromophenol blue (BPB) 1 % (v/v) stock

Per 10 ml

BPB	0.1 g
Tris-base	0.06 g

The lysis solution stock was made in a sterile universal and stored at room temperature.

2.5.2 Protocol for protein extraction from *E. coli* TOP10

Cells were harvest by centrifuging 40 ml of culture for 30 min at 4000 \times g at 4°C . The cell pellet was washed 3 times for 10 min at 4000 \times g at 4°C with 5 ml of 1 \times PBS. Then the pellet was resuspended in 100 μ l of lysis solution A. The cells were lysed by sonication on ice at an amplitude of 14 microns for 1 min in 10 s bursts with 10 s intervals. The protein mixture was harvested by centrifuging at 14000 \times g at 4°C for 45 min and the supernatant saved.

2.5.3 Separation of proteins using SDS-PAGE electrophoresis

2.5.3.1 SDS-PAGE reagents

12 % (w/v) acrylamide resolving gel components

40 % (w/v) solution (37.5:1 acrylamide: bisacrylamide)	3.0 ml
Soluion B *	2.5 ml
18.2 MΩ/cm H ₂ O	4.5 ml
10 % (w/v) ammonium persulphate (APS)	50 µl
TEMED	10 µl

* Solution B

Per 100 ml	
2 M Tris-HCl pH 8.8	75 ml
10 % (w/v) SDS	4 ml

4 % (w/v) acrylamide stacking gel components

40 % (w/v) solution (37.5:1 acrylamide: bisacrylamide)	0.5 ml
Soluion C *	1.0 ml
18.2 MΩ/cm H ₂ O	2.5 ml
10 % (w/v) APS	30 µl
TEMED	10 µl

* Solution C

Per 100 ml	
1 M Tris-HCl pH 6.8	50 ml
10 % (w/v) SDS	4 ml

SDS-PAGE running buffer 10 × (Stock)

Running buffer was made up at 10 × stock concentration and diluted 1:10 before use.

Per Litre

Tris- HCl	30.3g
Glycine	144.0g
SDS	10.0g

The pH of the buffer was adjusted to 8.8 in a volume of ~900 ml prior to the addition of SDS.

SDS-PAGE sample buffer

Per 10 ml

60 mM Tris-HCl pH 6.8	0.6 ml
10 % (w/v) SDS	2.0 ml
50 % (v/v) glycerol	5.0 ml
14.4 mM β -mercaptoethanol	0.5 ml
1 % (w/v) BPB	1.0 ml

Stored at -20° C in 0.5 ml aliquots

Solublising SDS-PAGE sample buffer

Per 10 ml

SDS-PAGE sample buffer	7.6 ml
Urea	2.4 g

Stored at 4° C.

Protein size standards

High molecular weight range: 36, 45, 55, 66, 84, 97, 116 and 205 kDa

Low molecular weight range: 6.6, 14.2, 20, 24, 29, 36, 45 and 66 kDa

Lyophilised standards were reconstituted with 100 μ l of 18.2 M Ω /cm H₂O to give a final concentration of ~2.0-3.5 mg/ml, and aliquoted into 4 μ l amounts and stored at -20° C. The proteins used to produce the appropriate sized bands are listed in Appendix A2. When required for use 8 μ l of SDS- PAGE loading buffer was mixed with the thawed standard and boiled for 3 min before running.

Coomassie blue gel stain solution

Per Litre

Coomassie Blue R-250	1.0 g
Glacial acetic acid	100.0 ml
Methanol	450.0 ml

Coomassie gel destain solution

Per Litre

Glacial acetic acid	100.0 ml
Methanol	100.0 ml

2.5.3.2 SDS-PAGE

SDS-PAGE gels were cast between two clean glass plates (Bio-Rad) with dimensions of 10.1 \times 8.2 cm. The plates were separated by a 1 mm spacer ridge on the larger of the two plates. The plates were clamped together and checked to ensure the bottoms edges of the plates were flush. The clamp unit was secured with vertical downward pressure in a casting stand on a rubber gasket. The components for the resolving gel were combined, mixed, and slowly pipetted into the space between the plates, until the solution was approximately 2 cm below the top of the smallest plate. To prevent

oxygen from inhibiting polymerisation and to produce a straight top edge, a small volume of 18.2 MΩ/cm H₂O was then added to the top of the solution. The resolving gel was then allowed to polymerise for ~45 min. Once the gel was set, the 18.2 MΩ/cm H₂O on top of the gel was removed. The components of the stacking gel were added, mixed, and pipetted up to the top of small plate. A 10-toothed comb (1.1 × 0.75 cm) was then immediately inserted into the gap between the plates, at an angle so as to prevent air bubbles from getting trapped between the comb and the stacking gel, and the stacking gel was then allowed to polymerise for ~ 30 min. The comb was then removed to reveal the wells, which were immediately rinsed with 18.2 MΩ/cm H₂O to remove any gel debris and unpolymerised acrylamide. The gel was then placed vertically into the buffer tank, at an angle so as to prevent air bubbles from getting trapped at the bottom of the plates, (Appendix B8) and secured in a holder with another gel, constructed in exactly the same manner on the opposite side. SDS-PAGE (1 ×) running buffer was then poured into the tank ensuring the level of buffer in the space between the two gels was higher than the level on the outside of the gels.

Loading samples were prepared for SDS-PAGE by the addition of 15 µl of 18.2 MΩ/cm H₂O to 5 µl samples, and then 5 µl of SDS-PAGE loading buffer was added. The samples were then pulsed briefly prior to boiling for 3 min. Size standards were prepared in the same way, except 8 µl of loading dye was added to 4 µl of size standard. All samples and standards were then centrifuged at 14,000 × g for 1 minute prior to loading 20 µl of all samples and 10 µl of all standards into the wells using a Hamilton syringe. The pair of gels was then electrophoresed at 120 milliamps, 200 volts for ~ 50 min, or until the BPB dye had started to come off the bottom of the gels.

2.5.3.3 Visualisation of protein bands and photography of SDS–PAGE gels

After electrophoresis, the gels were removed from the apparatus and stained for 10 min in Coomassie blue stain followed by immersion in destain overnight until crisp band images were visible against a clear background. The gels were then photographed using a gel documentation system (Bio-Rad Gel Doc 2000 with CCD camera using Quantity One™ software, Appendix B9). Hard copies of the gel picture were produced

using a Mitsubishi Video Copy Processor (Model P91 attached to the gel doc system), with Mitsubishi thermal paper (K65 HM-CE / High density type, 110 cm × 21 m).

2.5.4 Separation of proteins by 2DE

The protocol for 2DE followed 2DE using immobilized pH gradient—principles and methods (AmershamBiosciences, 2001).

2.5.4.1 Rehydration of Immobiline DryStrip gels

2.5.4.1.1 Rehydration solutions

Rehydration solution A stock

Per 10 ml

Urea	4.8 g
CHAPS	0.2 g
1 % (v/v) BPB	20 µl

The rehydration solution stock was divided into 700 µl aliquots and stored at – 20°C. Prior to use, 14 µl of IPG buffer and 1.4 mg of dithiothreitol (DTT) were added to each tube, giving a final concentration of 2 % (v/v) and 0.2 % (w/v) respectively.

2.5.4.1.2 Rehydration of Immobiline DryStrip pH 3-10, 18 cm with rehydration sample loading

To 80 µl of sample (proteins mixture in lysis solution A), 270 µl of rehydration solution A was added. A total of 350 µl of solution was therefore pipetted into one slot of the Immobiline DryStrip reswelling tray, which was levelled before use, and the sample was distributed along the slot. The protective cover was removed from the Immobiline DryStrip gel, starting at the acidic end (marked with a “+”). The gel strip

was positioned with the gel side down and the anodic (+) end of the strip oriented toward the pointed end of the tray. The Immobiline DryStrip gel was placed onto the solution, and was gently lifted and lowered, slid back and forth along the surface of the solution to remove all the bubbles between the DryStrip and the reswelling tray. Then 5 ml of Immobiline DryStrip Cover Fluid (Amersham Biosciences) was applied on top of DryStrip to minimize evaporation and prevent urea crystallization. The DryStrip gel was rehydrated at room temperature overnight.

2.5.4.1.3 Rehydration of Immobiline DryStrip pH 3-10, 18 cm with cup sample loading

The process of rehydration was the same as in section 2.5.4.1.2, except 350 μ l of rehydration solution A without sample was pipette into one slot of the Immobiline DryStrip reswelling tray.

2.5.4.2 First dimension - Isoelectric focusing (IEF)

IEF was performed using a Multiphor II electrophoresis system (Amersham Biosciences) incorporating an Immobiline DryStrip Kit which is listed in Appendix B10.

2.5.4.2.1 Preparation of the Immobiline DryStrip Kit

The cooling plate was placed on the Multiphor II system before positioning the Immobiline DryStrip Kit. Then 5 ml of Immobiline DryStrip Cover Fluid was pipetted on top of the cooling plate. The Immobiline DryStrip tray was placed on the cooling plate, and big bubbles were removed to ensure good thermal contact between the tray and the cooling plate. The red and black electrode leads on the tray were connected to the Multiphor II unit. Approximately 10 ml of Immobiline DryStrip Cover Fluid was poured into the Immobiline DryStrip tray, and the Immobiline DryStrip aligner was placed on top of the cover fluid, groove side up. Two electrode

strips were cut to a length of 11 cm each, and 500 μ l of 18.2 M Ω /cm H₂O was applied along each strip.

2.5.4.2.2 Preparation for IEF with rehydration loaded Immobiline DryStrips

The rehydrated DryStrip gel was removed from the reswelling tray and transferred in the groove of alinger in the Immobiline DryStrip tray, gel side up. The strip was placed with the acidic end at the top of the tray near the red electrode (anode), and another end at the bottom of the tray near the black electrode (cathode). The moistened electrode strips were placed across the cathodic and anodic ends of the aligned DryStrips, and contacted with the gel surface. The electrodes were aligned over the electrode strips, and the DryStrips were covered by 200 ml of cover fluid.

2.5.4.2.3 Preparation for IEF for cup loaded Immobiline DryStrips

For cup loading, after the Immobiline DryStrip Kit and electrode strips were prepared as in section 2.5.4.2.1, the sample cups and sample cup bar were placed at the anodic end of the apparatus. The sample cups were faced toward the electrode, and pressed down on the top of the Immobiline DryStrip gels. Good contact with the Immobiline Drystrip gels was necessary to avoid leakage of the samples. To check the leakage 70 ml of Immobiline DryStrip Cover Fluid was pour into the tray, and then an additional 150 ml Immobiline DryStrip Cover Fluid was added to completely cover the sample cup. The protein samples which were in 80 μ l of rehydration solution A were applied by pipetting under the surface of the Immobiline DryStrip Cover Fluid (Figure 2.2).

2.5.4.2.4 Running IEF for Immobiline DryStrip pH 3-10, 18cm

The cooling plate was kept at 20 °C using a Techne Circulator C-100 with TECAM Heat Exchanger 1000. The power was supplied by a EPS 3501 XL Power Supply (Amersham Biosciences) programmed according to Table 2.6.

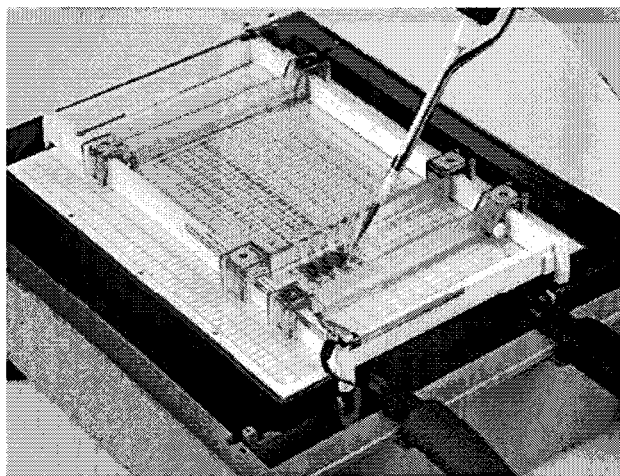


Figure 2.2 Applying sample into sample cups.

The figure was taken from Hand book: 2DE Using Immobilized pH Gradients, Principles and Methods (Amersham Biosciences)

After IEF was complete, the Immobiline DrySstrips were transferred to 25 × 200mm screw cap culture tubes. The second dimension electrophoresis was done immediately or the strips were store at -80 °C.

2.5.4.3 Equilibration of Immobiline Drystrip gels

Before the Immobiline Drystrip gels were applied to the second dimension electrophoresis apparatus, they were equilibrated.

2.5.4.3.1 Equilibration solutions

Equilibration buffer – 1.5 M Tris-HCl, pH 8.8

Per Litre

Tris base	181.7 g
-----------	---------

The pH of the buffer was adjusted to 8.8 with HCl. The solution was autoclaved and stored at room temperature.

SDS equilibration buffer solution stock

Per 200 ml

Equilibration buffer	10 ml
Urea	72 g
Glycerol	69 ml
SDS	2 g
1 % (v/v) BPB	200 µl

Step	Voltage (V)	Time (h)	KVh
1.Step and Hold	500	0:01	
2.Gradient	3500	1:30	3.0
3.Step and Hold	3500	4:50- 6:20	17- 22
Total		6:20- 7:50	20- 25

Table 2.6 Running conditions for Immobiline DryStrip pH 3-10, 18cm using a Multiphor II unit

The solution stock was divided into 20 ml aliquots in 28 ml universal, and store at -20 °C.

SDS equilibration buffer with DTT

Per 20 ml

SDS equilibration buffer solution stock	20 ml
---	-------

DTT	0.2 g
-----	-------

SDS equilibration buffer with iodoacetamide (IAA)

Per 20 ml

SDS equilibration buffer solution stock	20 ml
---	-------

IAA	0.9 g
-----	-------

2.5.4.3.2 Equilibration

Immobiline Drystrip gels were incubated at room temperature for 5 min. Then 10 ml of SDS equilibration buffer with DTT was added to each screw cap tube containing one Immobiline Drystrip gel. The tube was placed on the rocker to equilibrate for 15 min. Then the solution was poured off, and the same volume of SDS equilibration buffer with IAA was added to each tube. The gel was equilibrated in the same way for an additional 15 minute. After the equilibration was completed, the Immobiline DryStrip gel was transferred onto filter paper soaked with 18.2 MΩ/cm H₂O for 10 min.

2.5.4.4 Second dimension – SDS-PAGE

Second dimension – SDS-PAGE was performed using a Protean II XL 2-D cells (Bio-Rad). The major parts and accessories are listed in Appendix B11.

2.5.4.4.1 SDS-PAGE reagents

12 % (w/v) acrylamide resolving gel components

18.2 MΩ/cm H ₂ O	65.25 ml
1.5 M Tris- HCl, pH8.8	37.5 ml
10 % (w/v) SDS Stock	1.5 ml
40 % (w/v) solution (37.5:1 acrylamide: bisacrylamide)	45 ml
10 % (w/v) APS	750 µl
TEMED	75 µl

4 % (w/v) acrylamide stacking gel components

18.2 MΩ/cm H ₂ O	12.85 ml
0.5 M Tris- HCl pH6.8	37.5 ml
10 % (w/v) SDS Stock	200 µl
40 % (v/v) solution (37.5:1 acrylamide: bisacrylamide)	1.95 ml
10 % (w/v) APS	100 µl
TEMED	20 µl

SDS-PAGE running buffer 5 x (Stock)

Running buffer was made up at 5 x stock concentration and diluted 1:5 before use.

Per Litre

Tris base	15 g
Glycine	72 g

SDS

5 g

The buffer was kept at 4 °C for long-term storage. Before use, it was warmed up to 37 °C.

Protein size standards (Appendix A2)

Low molecular weight range (6.5, 14.2, 20, 24, 29, 36, 45, and 66 kDa) was used as protein size standards. The preparation was the same as in section 2.5.3.1.

2.5.4.4.2 Casting of SDS–PAGE gel

2.5.4.4.2.1 Assembling the glass plate sandwiches

SDS-PAGE gels were cast between two clean glass plates with dimensions of 20 x 20 cm and 20 x 22 cm. The plates were separated by two 1.5 mm spacers and an alignment card was used to set the spacers in place. The plates were clamped together with single-screw clamps and checked to ensure the bottom of the plates were flush and the clamp unit was secured with vertical downward pressure in a casting stand on a rubber gasket.

2.5.4.4.2.2 Casting the gels

The components for the resolving gel were combined, mixed and slowly pipetted into the space between the plates, until the solution was approximately 4 cm below the top of the smallest plate. To prevent oxygen from inhibiting polymerisation and to produce a straight top edge, 18.2 MΩ/cm H₂O was then added to the top of the solution. The resolving gel was then allowed to polymerise for ~1 h. Once the gel had set, the 18.2 MΩ/cm H₂O on top of the gel was removed using clean absorbent paper. The components of the stacking gel were added, mixed, and pipetted up to the top of small plate. A 1.5 mm thick comb with one large well of 17 cm and one small well for the

size standard was then immediately inserted into the gap between the plates at an angle to ensure that no air bubbles were formed underneath the well. The stacking gel was then allowed to polymerise for ~45 min. The comb was then removed to reveal the wells, which were immediately rinsed with 18.2 MΩ/cm H₂O to remove any gel debris and unpolymerised acrylamide.

2.5.4.4.3 Set up for SDS-PAGE

Agarose sealing solution

Per 20 ml

Running buffer (1 x)	19 ml
Agarose	0.1 g
1 % (v/v) BPB	40 µl

Agarose sealing solution were prepared by heating agarose in 1 × running buffer until the solid had completely dissolved, then 1µl of BPB was added to it.

The equilibrated Immobiline DryStrip gels were placed on the top of the stacking gel. The liquid agarose sealing solution was then pipetted into the space between two glass plates. The agarose sealing solution was left to cool down, sealing the strips in place.

After the Immobiline DryStrip gels were loaded on the SDS-polyacralamide gels, the gel sandwich was assembled on the cooling core. A upper tank was formed as the two gel sandwiches were attached on the cooling core, and 200 ml of 1 x running buffer was pour in it. On each gel on either side of cooling core 8 µl of protein size standards was loaded. It was then gently moved into the buffer tank. 2.5 L of 1 x running buffer was slowly pour in the tank to avoid any air bubbles produced at the bottom of each gel.

2.5.4.4.4 Running SDS-PAGE

The cooling core was kept at 10 °C using a Techne Circulator C-100 with TECAM Heat Exchanger 1000. The pair of gels was then electrophoresed at 80 mA, 100-400 volts for at least 4 h, until the BPB dye had started to come off the gel.

2.5.4.5 Visualisation and photography of 2DE gels

After electrophoresis, the 2DE gels were removed from the apparatus and visualized using Coomassie blue stain (section 2.5.3.1) following the protocol in section 2.5.3.3. The gels were then photographed using a gel documentation system, as detailed in section 2.5.3.3.

2.6 2DE of streptococci

2.6.1 Preparation of proteins from streptococci for 2DE

2.6.1.1 Solutions for protein extraction

Lysis solution B stock

Per 10 ml

Urea	4.8 g
CHAPS	0.4 g

The lysis solution stock was made, divided into 100 µl aliquots and store at – 20 ° C. Then 2 µl of IPG buffer was added just prior to use to make final concentration of 2 % (v/v).

Rehydration solution B stock

Per 10 ml

Urea	4.8 g
CHAPS	0.2 g
1 % (v/v) BPB	20 µl

The rehydration solution stock was made, divided into 700 µl aliquots and stored at – 20°C. Prior to use, 14 µl of IPG buffer and 2.8 mg of DTT were added to each tube, to give a final concentration of 2 % (v/v) and 0.4 % (w/v) respectively.

Other solutions

1x PBS, lysis solution A stock, rehydration solution A stock and 1 % (w/v) BPB stock were made as in section 2.5.1

2.6.1.2 Isolation of cell-associated proteins from *S. pyogenes* SF370

The protocol used for extraction of protein from *E. coli* TOP10 (section 2.5.2) was used to isolate the cell-associated proteins from *S. pyogenes* SF370. Proteins isolated from 40 ml of stationary phase cultures were separated by 2DE pH range 3-10, and IPG buffer pH 3-10 was used in the lysis solution and rehydration solution. Proteins isolated from 80 ml of cultures were separated by 2DE pH range 4-7 and 6-11. IPG buffers pH 4-7 and 6-11 was used, respectively.

2.6.1.3 Isolation of secreted proteins from *S. pyogenes* SF370 culture supernatant

S. pyogenes SF370 was centrifuged at 8,000 ×g, 4 °C for 15 min. Then supernatant (25 ml) was transferred into a clean tube, and precipitated by adding 1.5 ml of 100% trichloroacetic acid (TCA) to a final concentration of 10%. The tube was kept on ice for 10 min. The proteins were pelleted by centrifugation at 14,000 ×g at 4 °C for 15 min. The protein pellet was washed three times and centrifuged each time for 15 min at 14,000 ×g at 4 °C in 500 µl of cold acetone. Then the proteins were dissolved in 500 µl of cold 18.2 MΩ/cm H₂O. The process was then repeated from the beginning. After the

protein pellet was air dried, it was dissolved in 100 µl of lysis solution B, and was ready for further treatment.

2.6.1.4 Treatment with a 2-D Clean-Up kit

The protein extractions were treated with a 2-D Clean-Up kit (Amersham Biosciences, Appendix A3) to remove contaminants in the sample. The process followed the protocol provided with the product.

For each 100 µl of protein extraction, 300 µl of precipitant was added mixed by vortexing and incubated on the ice for 15 min. Then 300 µl of co-precipitant was added and mixed. Proteins were pelleted by centrifugation at 14,000 ×g, 4 °C for 5 min. The supernatant was removed and the remaining supernatant was brought down to the bottom of the tube by briefly centrifuging and removed the supernatant with a micropipette tip. Another 40 µl of co-precipitant was added on top of the protein pellet and incubated on the ice for another 5 min. The solution was removed and the proteins were resuspended in 25 µl of 18.2 MΩ/cm H₂O. Then 1 ml of pre-chilled wash buffer and 5 µl of wash additive were added into the protein suspension and incubated at -20 °C for 30 min. The suspension was vortexed for 20-30 s every 10 min. Then the tube was centrifuged at 14,000 ×g and 4 °C for 5 min to pellet the protein. The supernatant was removed and the pellet was air dried and dissolve in 100 µl of rehydration solution A or B for cup loading sample, or 350 µl of rehydration solution A or B for rehydration loading sample.

2.6.1.5 Determination of protein concentration

Before the proteins were separated by 2DE, the amount of protein present in the samples was determined.

2.6.1.5.1 Bradford method (Bradford, 1976)

2.6.1.5.1.1 Bradford assay reagents

BSA was made at a concentration of 100 mg/ml and was aliquoted into 1 ml amounts and stored at -20 ° C. Bradford reagent (Brilliant blue G 250 in phosphoric acid and methanol) was stored at 4 ° C.

2.6.1.5.1.2 Bradford assay

Protein determination by the Bradford method involved the construction of 14 standards of BSA in 5 µl aliquots ranging from 0.1- 1.4 mg/ml. Bradford's solution (250 µl) was then added to each dilution, and mixed on the shaker for 30 min. The mixtures were incubated at room temperature until they were analysed. Each standard was then measured in a 96-well plate at A_{570nm} against a blank containing 5 µl of 18.2 MΩ/cm H₂O mixed with 250 µl of Bradford's solution. All measurements were completed within 40 min of adding the Bradford's solution. A standard curve was then constructed (Figure D4), and the gradient of the initial linear part of the curve was used to calculate the X value for a given Y value using the equation for a straight line, $y = mx + c$ ($y = A_{570nm}$, x = the concentration of the sample proteins, m = the gradient of a straight line, and c = where the lines intercept with the y axis).

The protein samples and the dilutions of 1:2, 1:5 and 1:10 of protein samples in rehydration solutions were then measured after the addition of Bradford's solution. The absorbance values obtained were applied to the above equation to calculate the concentration of the proteins.

2.6.1.5.2 2-D Quant Kit (Appendix A4)

2.6.1.5.2.1 Reagents for 2-D Quant Kit (Amersham Biosciences).

BSA was made at a concentration of 2 mg/ml and was aliquoted into 1 ml amounts and stored at -20 ° C. The kit was maintained at 4 °C for long-term storage. Prior to performing the assay, colour reagent A and colour reagent B were mixed 100:1, respectively.

2.6.1.5.2.2 Assay using 2-D Quant Kit

Protein determination by 2-D Quant Kit involved the construction of 6 standards with 2 mg/ml BSA shown in Table 2.7. Also 5 µl and 10 µl of each sample were prepared in separately tubes. To each tube, 500 µl of precipitant was added, mixed by briefly vortexing and incubated at room temperature for 3 min. Then 500 µl of co-precipitant was added to each tube and mixed by inversion. Proteins were pelleted by centrifugation at 10,000 x g for 5 min. The supernatants were removed rapidly, and the tubes were centrifuged again to bring any remaining liquid to the bottom of the tube which was removed using a micropipette tip. Then the proteins were dissolved in 100 µl of copper solution and 400 µl of 18.2 MΩ/cm H₂O. After 1 ml of colour reagent A and colour reagent B mixture (section 2.6.1.6.2.1) was added, the tube was incubated at room temperature for 20 min. The standards and samples were measured at A_{480} against a blank containing 1 ml of 18.2 MΩ/cm H₂O. All measurements were completed within 40 min of initially adding the colour reagents. The standard curve was generated by plotting the absorbance of the standards against the quantity of protein, and the protein concentration of the samples was determined using the standard curve (Figure D5).

Tube No.	1	2	3	4	5	6
Volume of BSA standard (μl)	0	5	10	15	20	25
Protein quantity (μg)	0	10	20	30	40	50

Table 2.7 Preparation of standard curve for assay using 2-D Quant Kit

2.6.2 Separation of proteins from streptococci using 2DE

2.6.2.1 Rehydration of Immobiline DryStrip gels

For rehydration loading, the process of rehydration of Immobiline DryStrip gels followed the method detailed in section 2.5.4.1.2. The IPG buffer pH 3-10 and pH 4-7 were used in rehydration solution A/B, for rehydration of Immobiline DryStrip pH 3-10 and pH4-7, respectively.

For cup loading, Immobiline DryStrip gels were rehydrated using 350 µl of rehydration solution A/B, without protein samples, overnight at room temperature (section 2.5.4.1.3).

2.6.2.2 IEF for separation of proteins from streptococci

For IEF in pH range pH 4-7, rehydration sample loading was employed, and the process of IEF followed the method detailed in section 2.5.4.2.2. The programme used for focussing the Immobiline DryStrip pH 4-7, 18 cm was as Table 2.8.

For IEF in pH range 6 -11, see section 2.5.4.2.3.

The programme used for focussing the Immobiline DryStrip pH 6-11, 18 cm was as Table 2.9.

2.6.3 Equilibration of Immobiline DryStrip gels

The process of equilibration was followed as section 2.5.4.3.

Step	Voltage (V)	Time (h)	kVh
1. Step and Hold	500	0:01	
2. Gradient	3500	1:30	3.0
3. Step and Hold	3500	5:40-7:40	20-27
Total		7:10-9:10	23-30

Table 2.8 Focussing conditions for Immobiline DryStrip pH 4-7, 18cm using Multiphor II unit.

Step	Voltage (V)	Time (h)	kVh
1. Step and Hold	500	0:01	
2. Gradient	3500	1:30	3.0
3. Step and Hold	3500	5:40	20
Total		7:10	23

Table 2.9 Focussing conditions for Immobiline DryStrip pH 6-11, 18cm using Multiphor II unit

2.6.4 SDS-PAGE for separation of proteins from streptococci

2.6.4.1 SDS-PAGE reagents

14 % (w/v) acrylamide resolving gel components

18.2 MΩ/cm H ₂ O	50 ml
1.5 M Tris- HCl, pH8.8	37.5 ml
10 % (w/v) SDS Stock	1.5 ml
40 % (w/v) solution (37.5:1 acrylamide: bisacrylamide)	52.2 ml
10 % (w/v) APS	750 µl
TEMED	75 µl

Other reagents were made as section 2.5.4.4.1

2.6.4.2 Set up and running of SDS -PAGE

The 2nd dimension electrophoresis SDS-PAGE for separation of proteins from streptococci was performed as the method mentioned in section 2.5.4.4.

2.6.5 Visualisation of 2DE gels

2.6.5.1 Colloidal Coomassie Blue staining reagents

Colloidal Coomassie Blue stock

Per Litre	
Ammonium sulphate	100 g
Phosphoric acid	20 ml
Coomassie Blue G 250	1 g

Ammonium sulphate solution was made by dissolving 100 g in 20 ml of phosphoric acid. The Coomassie Blue G 250 was dissolved in 20 ml of 18.2 MΩ/cm H₂O. Then the two solutions were mixed and 18.2 MΩ/cm H₂O was added up to 1 litre. Colloidal Coomassie Blue stock was made at least 24 h in advance of staining, and was stored at room temperature.

Fix solution

Per Litre

Methanol	500 ml
----------	--------

Glacial acetic acid	120 ml
---------------------	--------

Staining solution

Four parts Coomassie stock was diluted to one part methanol. Before the stock was diluted, the stock was shaken to suspend the Coomassie stain.

2.6.5.2 Colloidal Coomassie Blue staining and destaining

After electrophoresis, the 2DE gels were removed from the apparatus and fixed using 250 ml of fix solution for at least 1 h. Then the gels were stained in 250 ml of staining solution overnight. The protein spots were visualized by washing the gels with 18.2 MΩ/cm H₂O.

2.7 Analysis of 2DE gels using PDQuestTM v 6.2.1 software

The analysis of 2-D electrophoresis gels included six steps which are shown in the Figure 2.3.

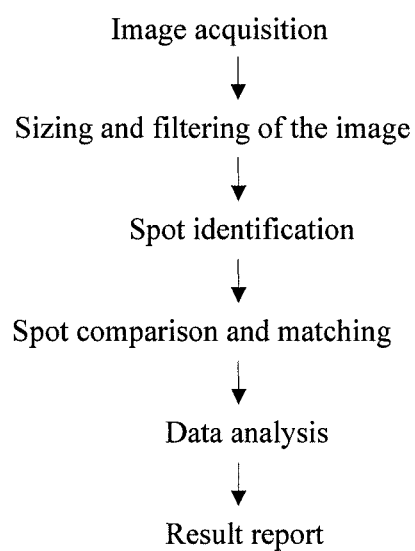


Figure 2.3 Process involved in analysis of 2-D electrophoresis gels using PDQuestTM v 6.2.1 software.

2.7.1 Image acquisition

The images of 2-D gels were acquired using Bio-Rad's GS-710 image densitometer through PDQuestTM v 6.2.1 software. Before the area of interest was scanned, the entire gel was preview by 'Preview Scan'. Then the area of interest which included all the protein spots were framed and the acquired. The signals from gels were converted into digital data. The data included the position of protein spots and the intensity of spots, which is shown as optical density (O.D).

2.7.2 Image sizing and filtering

2.7.2.1 Images sizing

When matching spots between the gels in a matchset, the matchset images should be all the same size and shape, therefore after 2-D gels images were acquired, all the images were cropped by the same setting which was done by selecting cropping area on one image, and landmarking a protein spot which was present in all the gels to be analysed.

2.7.2.2 Images filtering

Before the spots were identified by the software, small noise features, which would be able to affect the identification, were removed from the image using the filtering function. The noise characteristics were determined in the 'filter wizard'. The noise on the gels appeared as specks that were darker than surrounding background.

2.7.3 Detecting and editing spots

2.7.3.1 Detecting spots

The spot detection wizard was used to select the parameters for detecting the spots in gel scans. Because the spots on the gels had varying intensity, clarity, and streaking, different parameters were used to detect spots on gels that had been run and scanned at different times. The parameters were determined automatically after several features had been marked on the gel scan, which were faintest spot, smallest spot, largest spot and a region of the image including spots, streaks, background graduation, etc. Then the spots were detected based on this setting. When spots were detected in PDQuestTM v 6.2.1, the original gel scan was filtered and smoothed to clarify the spots, and then three-dimensional Gaussian spots were created from the clarified spots. Gaussian spots are 'ideal' spots that have been improved from fuzzy, streaked, or overlapping spots. The end result was three separate images: the original 2-D scan, which remained unchanged; the filtered image, which was a copy of the original scan that had been filtered and processed; and the Gaussian image, which was a synthetic image containing the Gaussian spots. The spot centres were marked by spot crosshairs, while spot shapes were marked by spot ellipses.

2.7.3.2 Editing spots

Several editions were performed on the spots after they had been detected automatically by the software.

2.7.3.2.1 Adding and removing spots

Although gels performed at the same time shared high similarity; they still differed. For example some significant spots had been missed on the gels which were not the one used to set up the parameters for spots detection. Adding spots function was used to add spots in the Gaussian image that had been missed out in the process of spots detection. For the same reason, some spots were incorrectly positioned or/and

misidentified. Removing spots function was used to remove those spots from the Gaussian image.

2.7.3.2.2 Combing spots and changing spot boundaries

Combing spots function was used to combine the quantitation of a cluster of spots into one spot. In addition, some spots were detected as saturated spots, which although appeared in the Gaussian image, were quantified were reported as zero. To quantify these spots appropriately, the spot boundary tools were used.

2.7.4 Spots comparison and matching

After the protein spots in the gels had been detected, the matchset was created to compare spots across gels. A matchset consisted of all gel images to be analysed, and the matchset standard, which was a synthetic image that contains the spot data from all the gels in that matchset.

2.7.4.1 Creating a matchset

A matchset was created from the Gaussian and Filtered images of all gels images, and the gel with the most spots was selected as the standard. When the matchset was created, standard spot (SSP) numbers were automatically assigned. PDQuestTM v 6.2.1 divides the standard into rows and columns, and in each section, there are approximately the same number of spots, but less than 1000 spots. The first two digits of a spot's SSP number correspond to its section's X and Y coordinates. Spots within a section were numbered sequentially.

2.7.4.2 Matching spots

Landmarking the key spots was the first step that matched all the spots in the matchset. Landmarks are reference spots used to align and position the gels for matching, and they were well-resolved and present in all members of the matchset. Using landmarks, PDQuestTM v 6.2.1 automatically matched most of the spots in the matchset. Then the remaining unmatched spots in the standard were reviewed one by one, and were either manually matched with existing spots in the standard, or were added to the standard if not present.

2.7.5 Data analysis

Three factors were considered during data analysis, and they were (i) the evaluation of quality; (ii) the determination of quantity and (iii) the determination of the quantitative changes in protein expression. The number of valid protein spots resolved in each gel and the number of spots matched to every gel were fundamental to the evaluation of the quality of the results. The scatter tool was used to show the quality of the matches.

When the gels in the same matchset were compared to each other, there was often some variation in spot size and intensity between gels that was not due to differential protein expression. These variation was caused by one or more factors, which are: (i) inconsistent cell numbers (O.D. of the culture) in sample preparation; (ii) inconsistent sample preparation efficiency due to variations in reagents; (iii) handling errors resulting in some ample loss during preparation; (iv) pipetting errors during sample preparation (v) sample loss during gel loading; and (vi) inconsistent staining times between gels. To accurately compare spot quantities between gels normalization was performed. Normalization in PDQuestTM v 6.2.1 was performed for each gel in each matchset using the following formula:

Normalized spot quantity = raw spot quantity * scaling factor/ normalization factor

Raw spot quantity is the unnormalized quantity (intensity) of each spot, and the scaling factor is a constant by which the normalized quantity is multiplied to give a more meaningful value, which was defaulted to 10^6 parts per million (ppm). The normalization factor was calculated for each gel based on ‘total quantity in valid spots’ method, i.e. the raw quantity of a particular spot in each member gel divided by the total raw quantity of all the spots in each member gel, that have been include in the standard. The method assumes that only a few protein spots change within the experiment. It is a useful method when very little information is known about the possible source of sample variation.

The analysis of quantitative alternations in protein expression was done on each gel. In the spot review page, the histogram of all spots on the standard is displayed, and each bar represents the spot’s quantity in each member of a matchset. Significant changes in protein expression were determined by the Wilcoxon two-sample test (Houtman et al., 2003). This was achieved by manually determining that the same change, either increase or decrease occurred in at least six sets of experiments, which is the minimum number necessary to achieve significance by Wilcoxon two-sample test. Finally all the spots that satisfied this criterion were created in a result set.

2.8 Protein digestion for MS analysis

2.8.1 Digestion of protein in solution

2.8.1.1 Reagents for digestion

Dissolve Solution

Per 100 ml

SDS	0.1 g
Tris base	6.06 g
DTT	77.1 mg

The pH was adjusted to 8 using HCl, and the solution was stored at room temperature.

Trypsin stock (1µg/µl)

The lyophilized powder of trypsin was stored at -20 °C.

Trypsin	100 µg
Glacial acetic acid 50 mM	100 µl

The trypsin powder was reconstituted in 50 mM glacial acetic acid and store at -20 °C for up to one month, and -70 °C for long-term storage.

2.8.1.2 Tryptic digestion

The concentration of protein was determined using the Bradford method (section 2.6.1.6.1). The proteins were added to 1 ml of 3.2 M ammonium sulphate, mixed well and pelleted by spinning at 14,000 xg for 5 min. To make a final protein concentration of 10 µg/µl, 20 µl of dissolve solution was added to the protein pellet. The proteins were dissolved by vortexing and heated to 60 °C for 50 min. After the solution was cooled down to room temperature, 5 µl of protein sample was run on a 12 % (w/v) SDS polyacrylamide gel as a control to check the extent of digestion. Then 1µg/µl of trypsin stock was added to the reaction, at a ratio of trypsin:protein of 1:70. The reaction was incubated at 37 °C for at least 1 h. Then 5µl of reaction solution was run on a 12 % (w/v) SDS polyacrylamide gel to check the extent of digestion. The remainder of reaction was kept on ice or in freezer at -20 °C for long-term storage.

2.8.2 In-gel protein digestion

2.8.2.1 Reagents for in-gel protein digestion

100 mM NH₄HCO₃

Per 10 ml

NH_4HCO_3 79 mg

50 mM NH_4HCO_3

Per 10 ml

100 mM NH_4HCO_3 5 ml

Trypsin solution (20 $\mu\text{g}/\text{ml}$)

Per 100 μl

1 $\mu\text{g}/\mu\text{l}$ trypsin stock 2 μl

50mM NH_4HCO_3 98 μl

The trypsin solution was made just prior to use.

2.8.2.2 In-gel digestion for colloidal Coomassie blue staining

Protein spots were cut into 1 mm x 1mm gel slices, and transferred to clean 1 ml silanised microfuge tubes. Colloidal Coomassie blue stain was removed by the addition of 100 μl of 100 mM NH_4HCO_3 and 60 μl of acetonitrile (ACN) in each tube. The tubes were shaken for 30 min and the liquid was removed. The wash step was repeated three times. Then the protein gels were dehydrated by addition of 50 μl of ACN, and incubated at room temperature for 5 min. Then the liquid was removed from the tubes, and the dehydration step repeated. The gel slices were dried in a centrifugal evaporator (Appendix B12) for 15 min, and were preincubated in 25 μl of 20 $\mu\text{g}/\text{ml}$ trypsin solution on ice for 30 min. Then 30 μl or more of 50 mM NH_4HCO_3 was added to each tube to cover the gel. The digestion was incubated at 37 °C overnight. The reaction was stopped by addition of 30 μl of 50 % (v/v) ACN/ 5 % (v/v) trifluoroacetic acid (TFA), and shaken for 30 min. The liquid was recovered and kept as it contains the digested peptides. Further digested peptides were recovered from the gel slices by adding 30 μl of 83 % (v/v) ACN/ 0.2 % (v/v) TFA, and shaken for 30 min. Again the liquid was saved and pooled with the previous extract.

2.8.3 Desalting using C₁₈ ZipTip® (Millipore)

2.8.3.1 Solutions for desalting

Wetting solution

Per 1 ml

ACN 500 µl

Elution solution

Per 1 ml

Methanol 500 µl

Glacial acetic acid 1 µl

Sample preparation, Wash, Equilibration solutions

Per 1 ml

TFA 1 µl

2.8.3.2 Desalting

The digests either from digestion in solution or in-gel was dried using a centrifugal evaporator at 60 °C for 30 min. 10 µl of sample solution was added to each dry digests. C₁₈ Ziptips were rehydrated by aspirating and dispensing 10 µl of wetting solution twice, and then were equilibrated by aspirating and dispensing 10 µl of equilibration solution twice. The peptides were bound onto the C₁₈ resin by aspirating and dispensing the samples up to 10 times. The salt was removed by aspirating the wash solution in and dispensing to waste, which was repeated twice. Elution solution (10 µl) was dispensed into a clean microfuge tube using a standard pipette tip. The peptides

were eluted by aspirating and dispensing elution solution through the ZipTip at least three times. Then the sample was ready for MS analysis.

2.9 MS analysis with LCQ Advantage (ThermoElectron)

2.9.1 Static nanospray MS set up

2.9.1.1 Static nanospray source

The static nanospray ion source included an NSI flange, and NSI probe assembly with static NSI probe and a probe cover (Figure 2.4).

The nanospray extension assembly and static NSI source was attached on the Finnigan LCQ Advantage MS detector.

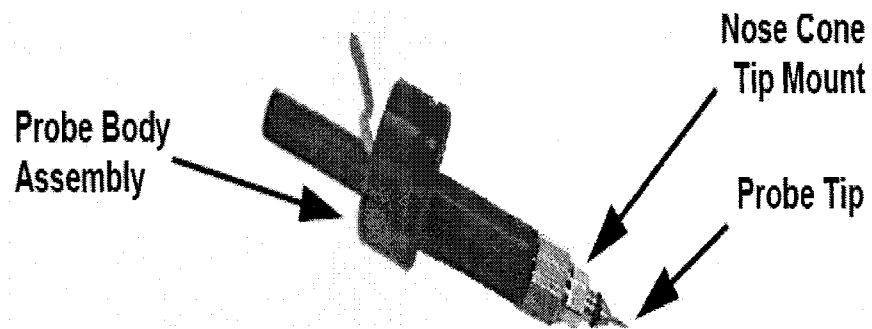
2.9.1.2 Installation of the static emitter

The standard coating EconoTip™ 1.2 mm / 1 µm (OD / tube ID) (New Objectives) was used as a sample emitter. The emitter was cut at the flat end to a length of 2.5 to 2.7 cm prior to mounting it on a static NSI probe. The emitter was inserted into the nose cone tip mount from the flat end of the emitter, and it was gently pushed further into the assembly until 1 to 1.5 mm of the emitter was left outside of the probe tip.

2.9.1.3 Loading sample into a static nanospray emitter

Before the emitters were installed on static NSI probe, the samples were loaded using a Gilson pipette fitted with a 20 µl GELoader tip (Eppendorf). The tips were inserted into the static nanospray emitter from the flat end after it was cut. 10 µl of each sample was injected slowly into the emitter to avoid the formation of air bubbles.

(a)



(b)

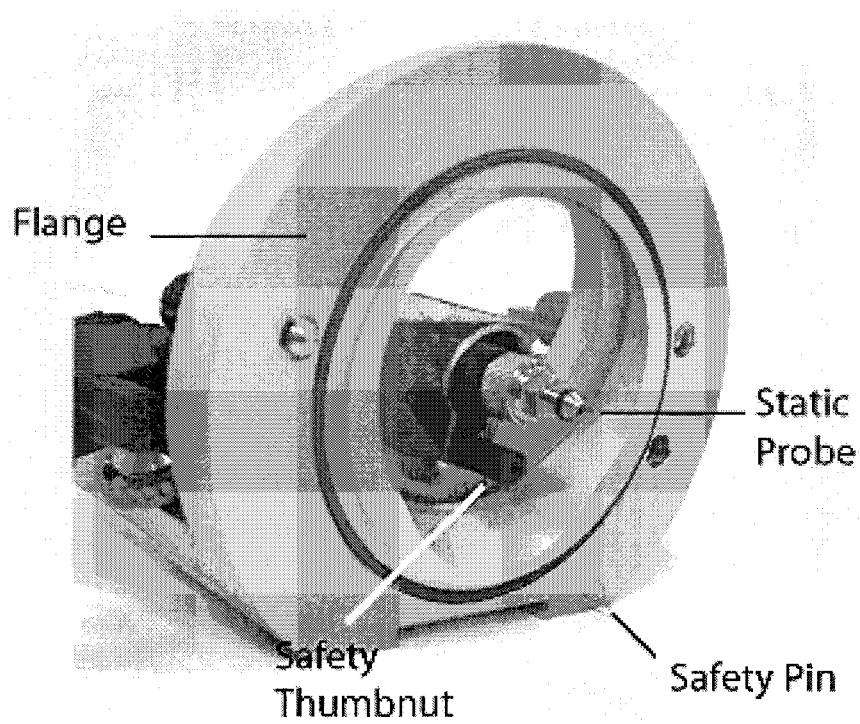


Figure 2.4 (a) Static NSI probe assembly. (b) Static NSI source

2.9.2 Tuning with the static NSI source

Tuning the Finnigan LCQ Advantage MS detector is essential for obtaining optimum performance. Tuning maximizes ion signal intensity.

2.9.2.1 Tuning solutions

All chemicals used for MS analysis are HPLC grade.

Angiotensin I (human) 200 μ M stock

To make a 200 μ M stock, 3.5 ml of 18.2 M Ω /cm H₂O was added to a vial of 1 mg Angiotensin I. The stock was stored at -20 °C.

Angiotensin I 1 μ M

Per 1 ml

Angiotensin I (human) 200 μ M stock	5 μ l
Methanol	500 μ l
Glacial acetic acid	1 μ l

The solution was made freshly prior to use, and kept on ice.

2.9.2.2 Tuning

10 μ l of 1 μ M Angiotensin I was loaded into an emitter and the static nanospray source was set up as section 2.9.1, and the distance from the spray tip to the ion transfer capillary was adjusted to \sim 2 mm. Before the instrument was turned on, NSI was selected as the API source type in the Tune Plus window, the spray voltage was set to 1.0 kV and the sheath gas and aux gas were set to 0. Then the instrument was turned

on. The position of nanospray tip and values for spray voltage and capillary temperature were adjusted until signal at m/z 1296.7 (singly charged), 648.8 (doubly charged) and 432.9 (triply charged) became stabilized. Then the tuning process was started by selecting automatic tab and input m/z 648.8 as selected mass. The LCQ Advantage MS detector needs to be tuned every week, and the result was saved as 'static NSI tune method'.

2.9.3 Analysis of salts and detergent tolerance of the static NSI source

2.9.3.1 Preparation of protein digests for analysis

Elution solutions for analysis of salts tolerance

A 1.6 M NaCl stock made in 50 % (v/v) methanol / 0.1 % (v/v) acetic acid, was serially diluted with the same solution to get elution solutions containing 1.6×10^{-1} , 1.6×10^{-2} , 1.6×10^{-3} and 1.6×10^{-4} M NaCl.

Elution solutions for analysis of detergent tolerance

A 1 M SDS stock was made by addition of 2.88 g in 10 ml of 50 % (v/v) methanol / 0.1 % (v/v) acetic acid was serially diluted with the same solution to produce solutions containing 1×10^{-1} , 1×10^{-2} , 1×10^{-3} , 1×10^{-4} and 1×10^{-5} M SDS.

Protein digests

Nine proteins (β -Galactosidase, phosphorylase b, albumin, L-glutamic dehydrogenase, ovalbumin, glyceraldehydes-3-phosphate dehydrogenase, carbonic anhydrase, trypsinogen and α -lactalbumin) were trypsin digested follow the method in section 2.8.1, and desalted using C_{18} ZipTips (section 2.8.3). The peptides was eluted using 50 % (v/v) methanol / 0.1 % (v/v) acetic acid containing 1.6×10^{-1} , 1.6×10^{-2} , 1.6×10^{-3} and 1.6×10^{-4} M NaCl or containing 1×10^{-2} , 1×10^{-3} , 1×10^{-4} and 1×10^{-5} M SDS.

Totally, 12 samples (A-L) were made for analysis of salts tolerance and 12 samples (a-l) were made for analysis of detergent tolerance (Table 2.10).

2.9.3.2 Collecting data

10 µl of each sample was loaded into the emitters, and the static nanospray source was set up as section 2.9.1. The 'static NSI tune method' was selected through Tune Plus window. Full scan type and scan mass range m/z 150-2000 were selected through Define Scan dialog. Through Acquire Data dialog, 100 scans for each sample were collected.

2.9.4 MS analysis for protein identification using the static nanospray

Data for protein identification was collected through XcaliburTM v 1.3 software (ThermoElectron).

Sample	Protein (μg)	NaCl/SDS concentration (M)
A	50	1.6×10^{-4}
B	50	1.6×10^{-3}
C	50	1.6×10^{-2}
D	50	1.6×10^{-1}
E	5	1.6×10^{-4}
F	5	1.6×10^{-3}
G	5	1.6×10^{-2}
H	5	1.6×10^{-1}
I	0.5	1.6×10^{-4}
J	0.5	1.6×10^{-3}
K	0.5	1.6×10^{-2}
L	0.5	1.6×10^{-1}
a	50	1×10^{-5}
b	50	1×10^{-4}
c	50	1×10^{-3}
d	50	1×10^{-2}
e	5	1×10^{-5}
f	5	1×10^{-4}
g	5	1×10^{-3}
h	5	1×10^{-2}
i	0.5	1×10^{-5}
j	0.5	1×10^{-4}
k	0.5	1×10^{-3}
l	0.5	1×10^{-2}

Table 2.10 Samples for analysis of salts and detergent tolerance.

2.9.4.1 Defining an instrument method

An instrument method was defined through Xcalibur™ v 1.3 software. The parameters were set up on the general MS or MS/MS function as follows:

Run settings

Acquired Time:	10 min
Segments:	1

Segment 1 Information

Duration:	10 min
Number of Scan Event:	3
Tune Method:	Static NSI

Scan Event Details

1: Positive	Full scan (350 -2000)
2: Positive	Zoom most intense ion from (1)
3: Positive	MS/MS most intense ion from (1)

Data Dependent Settings

Reject Mass List:	195.00, 279.00, 370.90, and 444.90*
Default Charge State:	2
Default Isolation Width:	2.00
Normalized Collision Energy:	35
Activation Q:	0.250
Activation Time:	30.000
Min. Signal Required	100000
Min. MSn Signal Required	5000

Charge state screening enable

Threshold (%): 2.0

* The reject masses was measured by inject 50 % (v/v) methanol / 0.1 % (v/v) acetic acid into MS detector. The peaks with most intensity were recorded.

Global Data Dependent Settings

Exclusion mass width by mass

Exclusion mass width low: 1.50

Exclusion mass width high: 1.50

Reject mass width by mass

Reject mass width low: 0.50

Reject mass width high: 1.50

ZoomScan mass width by mass

ZoomScan mass width low: 5.00

ZoomScan mass width high: 5.00

Dynamic exclusion enabled

Repeat Count: 1

Repeat Duration: 0.50

Exclusion List Size: 50

Exclusion Duration: 3.00

Exclusion mass width by mass

Exclusion mass width low: 1.50

Exclusion mass width high: 1.50

After the above parameters were set, the instrument method was saved as the 'static NSI method'.

2.9.4.2 Acquiring data using the static nanospray

The spectra of each protein digests was preview in Tune Plus window. The 'Static NSI tune method' was selected, and the instrument was turned on. When the signals were stable, the Tune Plus window was closed with the instrument on. The data was acquired through 'Run a Sample' function in Xcalibur v 1.3. The 'static NSI method' was selected as the instrument method, and LCQ Advantage MS was selected as 'In Use and Start Instrument'. As per the method set up, each run was collected for 10 min, and the process was monitored through 'Real Time Plot' function.

2.10 Protein analysis using BioWorks v 3.2 software (ThermoElectron)

TurboSEQUEST programme in BioWorks v.3.2 software was used to identify the protein or proteins present in a sample by matching the MS/MS spectra from the sample with theoretical MS/MS spectra generated from protein or nucleotide sequence databases. To obtain the identity of proteins, the spectra of protein digests acquired using the method defined in section 2.9.4 were opened with TurboSEQUEST and searched against protein or nucleotide sequence databases. There were three steps: creating an appropriate database, performing a TurboSEQUEST search and filtering the match list.

2.10.1 Creating a streptococci database

FASTA Database Utilities dialog was used to create a streptococci database from a larger database 'nr.fasta', which was supplied by Thermo Electron. 'nr.fasta' was selected as the original database, and 'streptococci.fasta' was defined as the target database. 'Streptococcus' was entered in the select parameters for inclusion in the database, then the new database was built as all the entries included 'streptococcus'.

'streptococci.fasta' database was updated regularly by adding new entries to the database, which was also achieved using the FASTA Database Utilities dialog. The completely sequenced genomes of streptococcal species including *S. agalactiae*

NEM316, *S. pyogenes* MGAS8232, *S. agalactiae* A909, *S. pyogenes* MGAS5005, *S. pyogenes* MGAS6180, *S. thermophilus* CNRZ1066, *S. thermophilus* LMG 18311, *S. pyogenes* MGAS10394, *S. mutans* UA159, *S. agalactiae* 2603V/R, *S. pyogenes* SSI-1, *S. pyogenes* MGAS315, *S. pneumoniae* R6, *S. pneumoniae* TIGR4, and *S. pyogenes* M1 GAS were downloaded from NCBI website (<http://www.ncbi.nih.gov/>). The sequences were pasted in the Enter box of the Adding function, and new entries were added by pressing Add Entry button.

2.10.2 Performing TurboSEQUEST searches

The TurboSEQUEST Search parameters were set as follows:

Perform DTA Generation

Scan limits:	1- 300
Instrument type:	LCQ
Use Charge state:	Auto
MSn lever:	Auto

Dta Generation Parameters for LCQ

MW range:	300 – 2000
Threshold:	Absolute 50000
Precursor ion tolerance:	1.2000
Mass Units:	AMU
Group scan:	0
Minimum ion count	25

Perform Search

C:\Xcalibur\params\streptococci.params

Database:	streptococci (FASTA)
Type:	Protein
Mass type:	Average precursor / Monoisotopic fragments
Enzyme:	Trypsin (KR)
Enzyme limits:	Fully enzymatic – cleaves at both ends
Missed cleavage sites:	2
Peptide tolerance:	1.0000
Units:	AMU
Fragment ions tolerance:	1.0000
Number results scored:	250

The raw data was searched against the 'streptococci.fasta' database using the parameters above, and the matched peptide and protein sequences were shown in the MultiConsensus Results window.

2.10.3 Filtering the match list

In the MultiConsensus Results window, the results table shows the reference of peptides including peptide sequences and charges (z), probability calculations, cross correlation score ($Xcorr$), difference in normalised cross correlation score (ΔC_n) and other values for the matched peptides and proteins from the TurboSEQUEST search. The proteins are ranked according to the matched peptide probability value. To select the positive results, several filtering parameters were set up: proteins must be from the SF370 strain of the M1 serotype of GAS; and peptides must have a $Xcorr > 1.5$ for singly charged peptides, > 2.0 for doubly charged peptides, and > 2.5 for triply charged peptides, with at least 2 peptides successfully matched; and/or must have a $\Delta C_n > 0.2$ with at least 2 peptides successfully matched (Eng et al., 1994). The best hit that fulfilled the filter parameters was selected.

2.11 Protein identification and characterisation using TSSViewTM v1.0 (Turbinia/Biosystems Informatics Institute)

TSSviewTM v1.0 was used to identify regions of evolutionary conservation in proteins as determined by high-sensitivity data-mining conducted on approximately three million UniProt entries and their corresponding coding and non-coding regions in each of the six reading frames.

2.11.1 Protein identification via functional conserved sequence

The proteins which produced peptides that satisfied the Xcorr and/or ΔCn values but did not satisfy the other criteria were analysed using TSSviewTM v 1.0 (including those peptides not from the SF370 strain of the M1 serotype of GAS). Observed / Expected (O/E) values $\leq 1.0 \times 10^{-4}$ were considered significant with the level of probability varying according to signature peptide length.

The proteins from the SF370 strain of the M1 serotype of GAS that contained a single highly-conserved peptide sequence were included in the list of identified proteins. For those proteins not from SF370 strain of the M1 serotype of GAS, but containing more than one highly-conserved peptide sequence, a BlastPAM30 search of the conserved peptides was employed. If the highly-identical peptides were present in the same protein from the SF370 strain of the M1 GAS serotype, they were included in the list of identified proteins.

2.11.2 Functional characterisation of identified proteins

The full length sequences of all identified proteins were analysed by TSSviewTM v 1.0. O/E values $\leq 1.0 \times 10^{-4}$ were considered as highly significant with the level of probability varying according to signature peptide length. The function of the highly-conserved regions was assigned as the potential function of that protein.

3 Results for development of technology for proteomics using 2DE with MS

3.1 Background

To develop the basic 2DE techniques, *E. coli* TOP10 was used as a model bacterium. The development of 2DE techniques was divided into two main steps. Firstly, the cell-associated proteins were extracted and visualized by one dimensional 12 % (w/v) SDS-PAGE. Secondly, the cell-associated proteins were separated by 2DE. In this process, the basic set up for 2DE was developed using a Multiphor II electrophoresis system and Protean II XL 2-D cells. Gel visualisation was performed using Bio-Rad's Gel Doc 2000 system. Several experiments were carried out to optimise the conditions for separating protein extracts from *S. pyogenes* SF370, including using different amounts of cells, different solutions for sample preparation and different 2DE processes, e.g. using several staining methods for visualization of the protein spots, and using different pH range of Immobiline DryStrips to get better images for analysis. Also methods for extraction of proteins from culture supernatant from *S. pyogenes* SF370

In addition, the basic techniques of static nanospray LCQ Advantage MS were developed and operating parameters were optimised in this study. As salt and detergent tolerance are general shortcomings of electrospray MS (Glish and Vachet, 2003), the salt and detergent tolerances of static nanospray LCQ Advantage MS was determined. The parameters of BioWorks 3.2 software were optimized and proteins were identified.

3.2 Development of 2DE

3.2.1 2DE of *E. coli* TOP10

3.2.1.1 Extraction of cell-associated proteins from *E. coli* TOP10

Cell-associated proteins of *E. coli* TOP10 were isolated according to the method in section 2.5.2, and lysis solution A was used. The isolation of proteins was visualised on a SDS-PAGE gel (Figure 3.1).

3.2.1.2 Experiments to optimize conditions for basic 2DE

Cell-associated proteins from 80 µl extract of *E. coli* TOP10 stationary phase culture were separated using 2DE. Four experiments were set up to optimise the techniques. The common conditions for these experiments were: (a) 40 ml of cell culture was used to extract the proteins; (b) rehydration solution A was used to rehydrate the Immobiline DryStrips; (c) pH range 3-10, 18cm, Immobiline DryStrips were used and ran at conditions stated in section 2.5.4.2.4; (d) 12 % (w/v) of SDS-PAGE was used in the second dimension of electrophoresis; (e) the gels were visualized using Coomassie blue stain; (f) the images were acquired using the CCD camera of the Gel Doc 2000 system. The different conditions are as following:

Experiment No. 1 (Figure 3.2): (a) the protein sample was loaded by rehydration sample loading; (b) fresh Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF.

Experiment No. 2 (Figure 3.3): (a) the protein sample was loaded by cup sample loading; (b) fresh Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF.

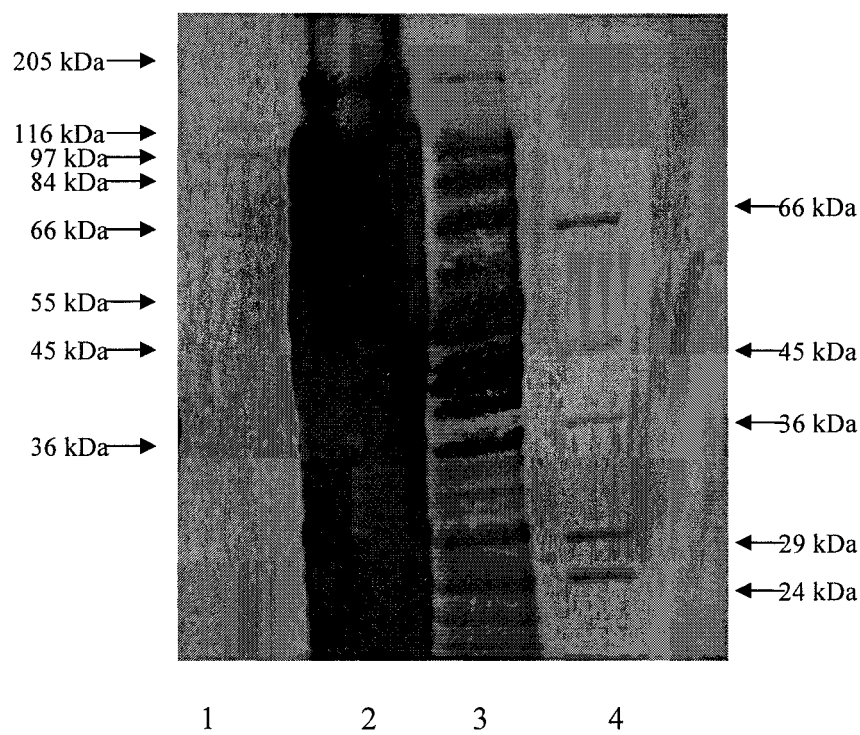


Figure 3.1 Visualization of cell- associated proteins from *E. coli* on 12 % (w/v) SDS-PAGE.

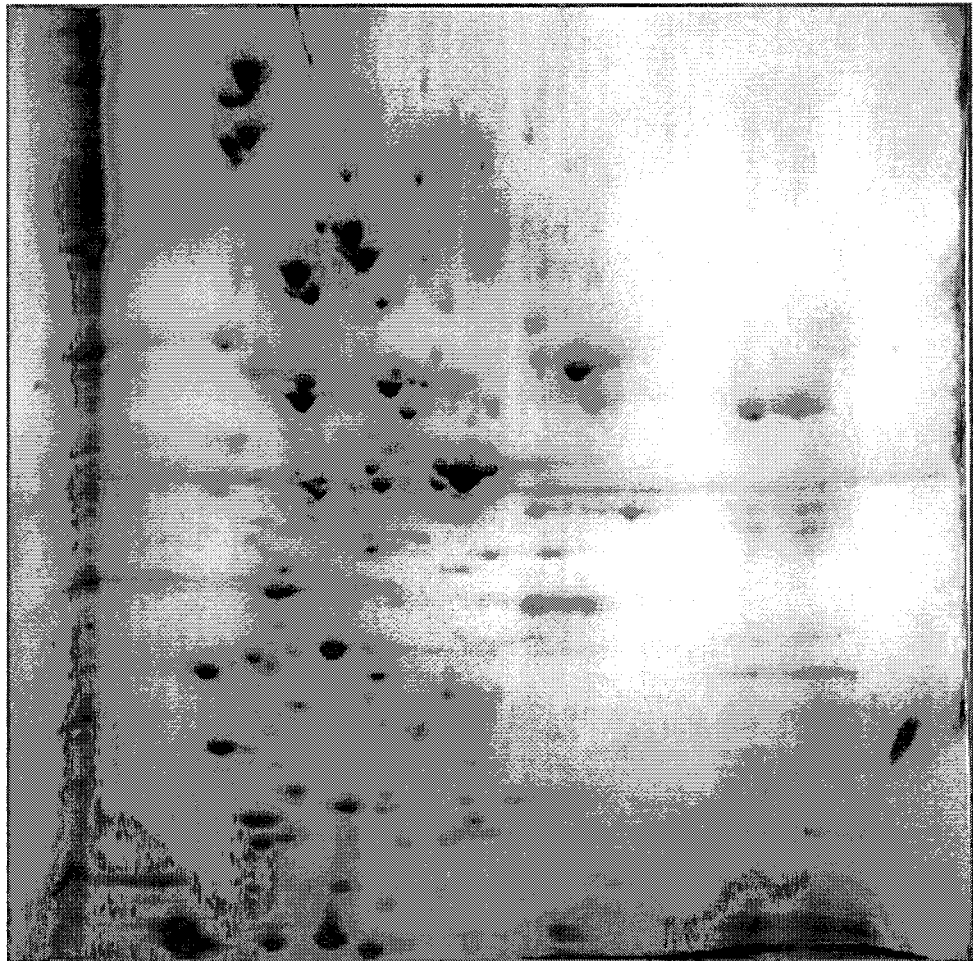
Lysis solution A was used to lyse the cells and extract the proteins. Lane 1: High molecular weight (HMW) makers: (M.W. 36, 45, 55, 66, 84, 97, 116 and 205 kDa); Lane 2: 5 μ l of original sample (section 2.5.2); Lane 3: 5 μ l of 10 fold dilution of original sample; Lane 4 Low molecular weight (LMW) markers (M.W. 24, 29, 36, 45 and 66 kDa).

Experiment No. 3 (Figure 3.4): (a) the protein sample was loaded by rehydration sample loading; (b) mineral oil was used as a replacement for Immobiline DryStrip Cover Fluid to cover the Immobiline DryStrips during IEF.

Experiment No. 4 (Figure 3.5): (a) the protein sample was loaded by rehydration sample loading; (b) recycled Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF.

These 2DE gels were compared (Figure 3.2 – 3.5) and it was obvious that the conditions used in experiment No.1 were optimal for 2DE in pH range 3-10 for separation of proteins from *E. coli*, as the other conditions produced gels which were of poor quality, i.e., there were only a few protein spots on the 2DE gel from experiment No.2; extreme horizontal streaking but no obvious spots were present on the 2DE gel from experiment No.3; and extremely horizontal streaking with only a few spots present on the 2DE gel from experiment No.4. Therefore, the conditions used in experiment No.1 were employed at the beginning of the studies of 2DE for separation of streptococcal proteins; i.e.:

- 40 ml of stationary phase culture was used to extract the proteins
- Lysis solution A was used in sample preparation
- Rehydration solution A was used in rehydration step
- Protein sample was loaded by rehydration sample loading when pH 3-10, 18cm Immobiline DryStrips were used for first dimension
- The fresh commercial Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF
- 12 % (w/v) SDS-PAGE was used in the second dimension of electrophoresis
- Coomassie blue stain was used to visualize the gels

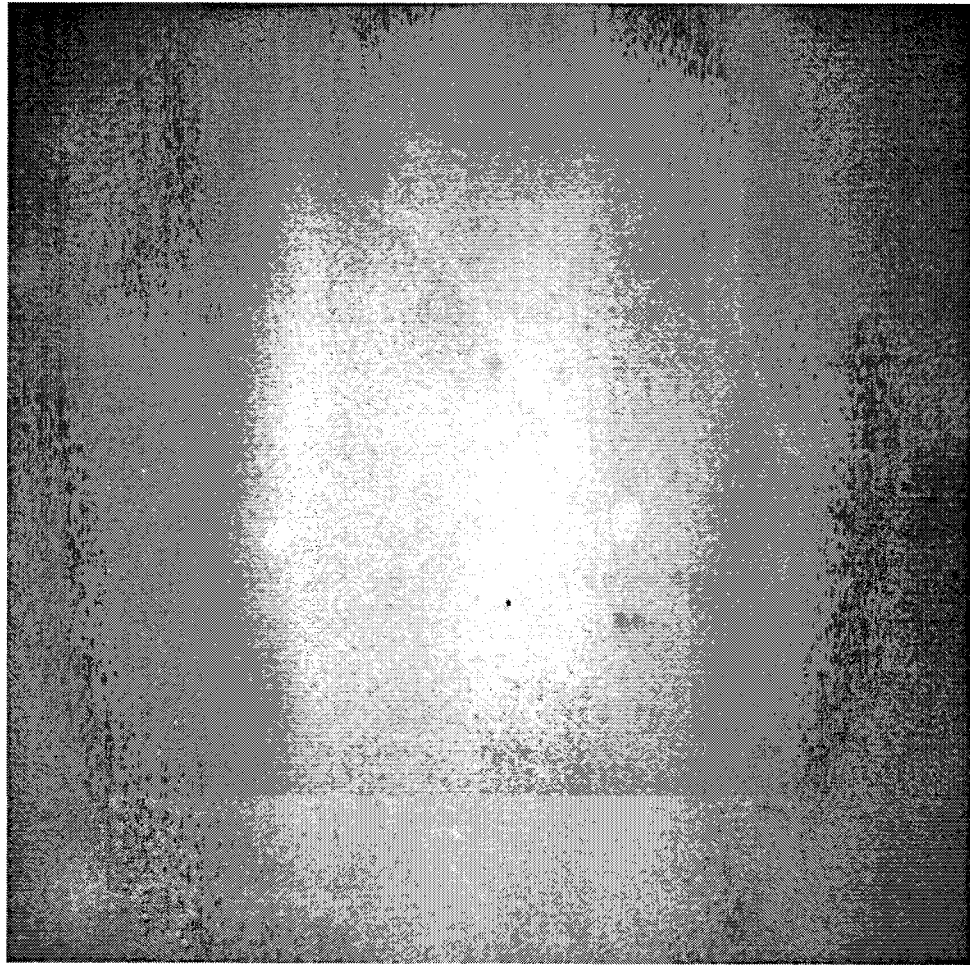


pH 3

10

Figure 3.2: 2DE gel of cell –associated proteins from *E. coli* (Experiment No.1)

(a) Protein sample was loaded by rehydration sample loading; (b) Fresh commercial Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF

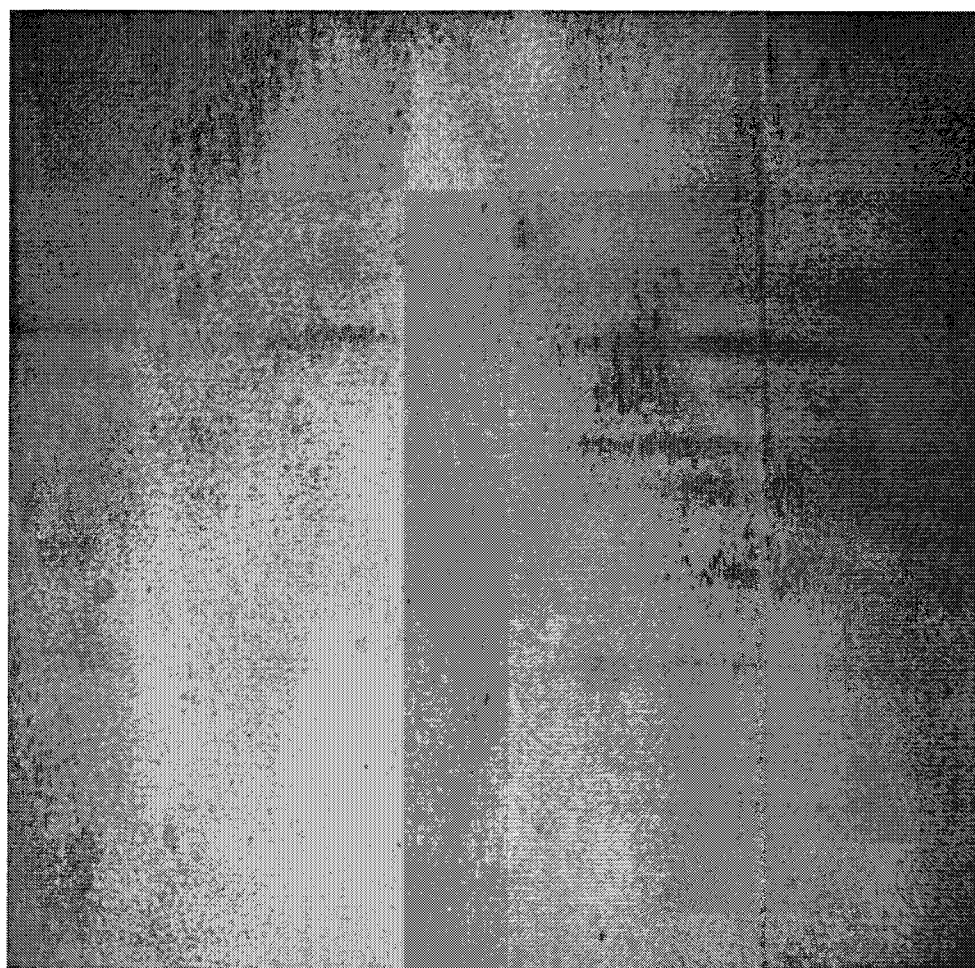


pH 3

10

Figure 3.3: 2DE gel of cell –associated proteins from *E. coli* (Experiment No.2)

(a) Protein sample was loaded by cup sample loading; (b) Fresh commercial Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF



pH 3

10

Figure 3.4: 2DE gel of cell –associated proteins from *E. coli* (Experiment No.3)

(a) Protein sample was loaded by rehydration sample loading; (b) Mineral oil was used as a placement of Immobiline DryStrip Cover Fluid to cover the Immobiline DryStrips during IEF



pH 3

10

Figure 3.5: 2DE gel of cell –associated proteins from *E. coli* (Experiment No.4)

(a) Protein sample was loaded by rehydration sample loading; (b) Recycled Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF.

3.2.2 2DE of streptococci

3.2.2.1 Development of 2DE for separation of proteins from *S. pyogenes* SF370

3.2.2.1.1 Separation of cell-associated proteins in pH range 3 -10

Immobiline DryStrips pH 3 -10, 18 cm were employed to obtain an overall view of the separation of cell-associated proteins from *S. pyogenes* SF370. Six experiments were carried out to optimize the techniques. The common conditions in these experiments were (a) Immobililine DryStrips pH 3 -10, 18 cm were employed; (b) 40 ml of cell culture was used to extract the proteins; (c) protein samples were loaded by rehydration loading; (d) the fresh Immobililine DryStrip Cover Fluid was used to cover the Immobililine DryStrips during IEF; (e) 12 % (w/v) SDS-PAGE was used in the second dimension of electrophoresis; (f) the images were acquired using the CCD camera of the Gel Doc 2000 system. The different conditions were as following:

Experiment No. 5 (Figure 3.6)*: (a) Lysis solution A was used to extract the proteins from cell lysate; (b) Rehydration solution A was used in the rehydration step; (c) Coomassie blue stain is used to visualize the gels.

Experiment No.6 (Figure 3.7)*: (a) Lysis solution B was used to extract the proteins from cell lysate; (b) Rehydration solution A was used in the rehydration step; (c) Coomassie blue stain was used to visualize the gels.

Experiment No. 7 (Figure 3.8)*: (a) Lysis solution B was used to extract the proteins from cell lysate; (b) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (c) Rehydration solution A was used in the rehydration step; (d) Coomassie blue stain was used to visualize the gels.

* The replicates of the 2D gels are shown in Figure E1.

Experiment No. 8 (Figure 3.9)*: (a) Lysis solution B was used to extract the proteins from cell lysate; (b) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (c) Rehydration solution A was used in the rehydration step; (d) colloidal Coomassie blue staining was used to visualize the gel.

Experiment No. 9 (Figure 3.10)*: (a) Lysis solution B was used to extract the proteins from cell lysate; (b) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (c) Rehydration solution B was used in the rehydration step; (d) colloidal Coomassie blue staining was used to visualize the gel.

The basic conditions for 2DE mentioned in section 3.2.1.2 were used in experiment No. 5. There was only a few distinct spot visible on the 2DE gel (Figure 3.6). The possible cause was insufficient proteins in the sample as low concentration of CHAPS can cause insufficient protein solubilisation. In experiment No. 6 lysis solution B was used, in which the CHAPS concentration was increased. There were more proteins present on the 2DE gel (Figure 3.7). However, horizontal streaking was also present on the gel. Non-protein impurities in the sample can interfere with IEF, causing horizontal streaking in the final result, particularly at the acidic end of the gel. Therefore in Experiment No.7, 2D clean up kit was employed to remove contaminants in the sample. On the resulting gel (Figure 3.4), there was still some streaking, and at the high molecular area of the gel, some protein spots appeared unclear. In experiment No. 8, colloidal Coomassie blue staining was used to visualize the gel. The gel images (Figure 3.8, and 3.9) were compared, and it was obvious that colloidal Coomassie blue staining was more sensitive than Coomassie blue staining. Therefore, colloidal Coomassie blue staining was employed throughout the project. However, the same problems were present on this gel, i.e. some streaking, and at the high molecular area, some protein appears unclear. In experiment No. 9, DTT, which reduces disulfide bonds, was doubled in concentration to allow proteins to unfold completely, which resulted in better solubilisation and also helped 'focus' the proteins that are present on the gel as unclear spots (Figure 3.10).

* The replicates of the 2D gels are shown in Figure E1.

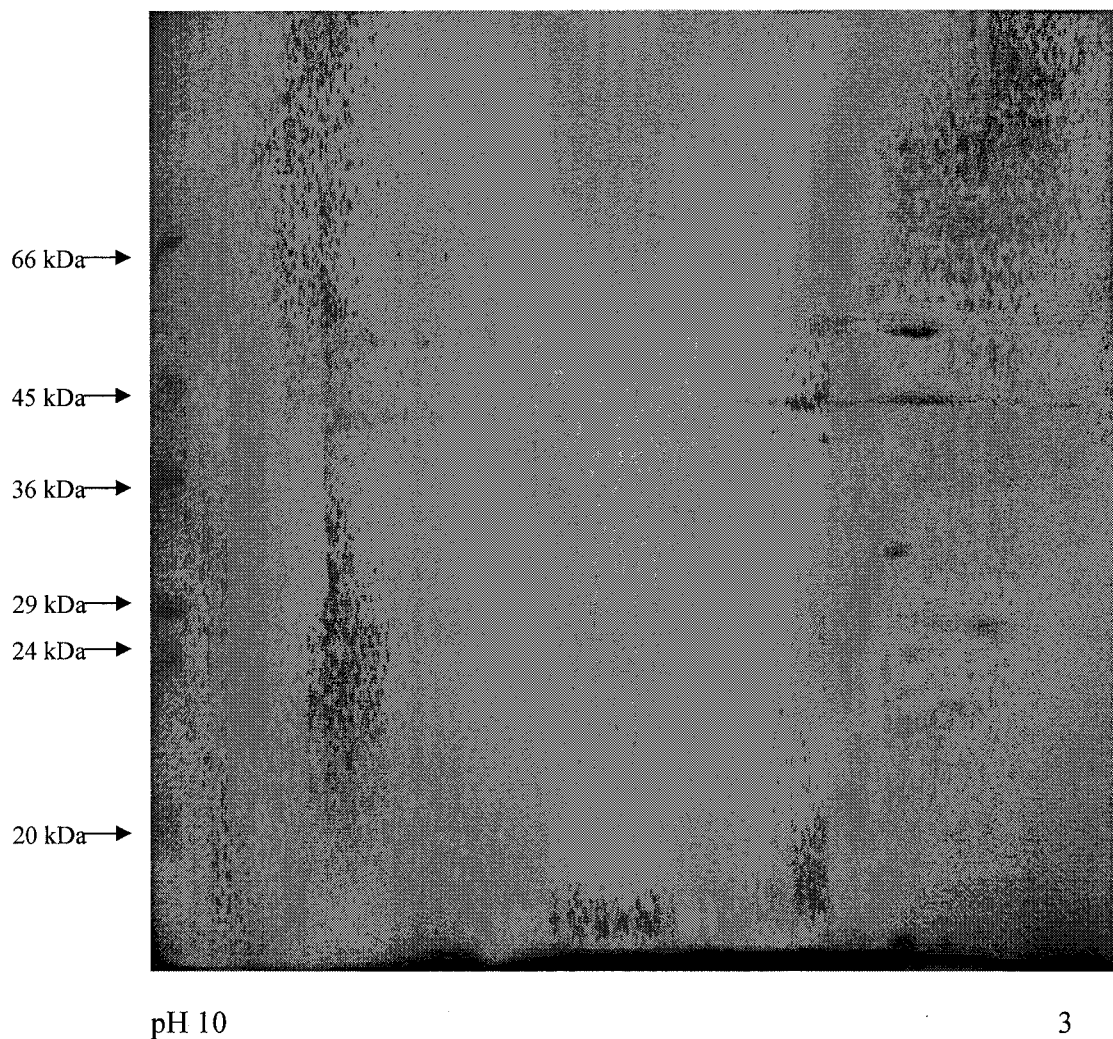


Figure 3.6: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.5).

(a) Immobiline DryStrips pH 3 -10, 18 cm were employed for IEF, and 12 % (w/v) SDS-PAGE was used for the second dimension; (b) Lysis solution A was used to extract the proteins from cell lysate; (c) Rehydration solution A was used in the rehydration step; (d) Coomassie blue stain was used to visualize the gel; (f) LMW markers: 66, 45, 36, 29, 24 and 20 kDa.

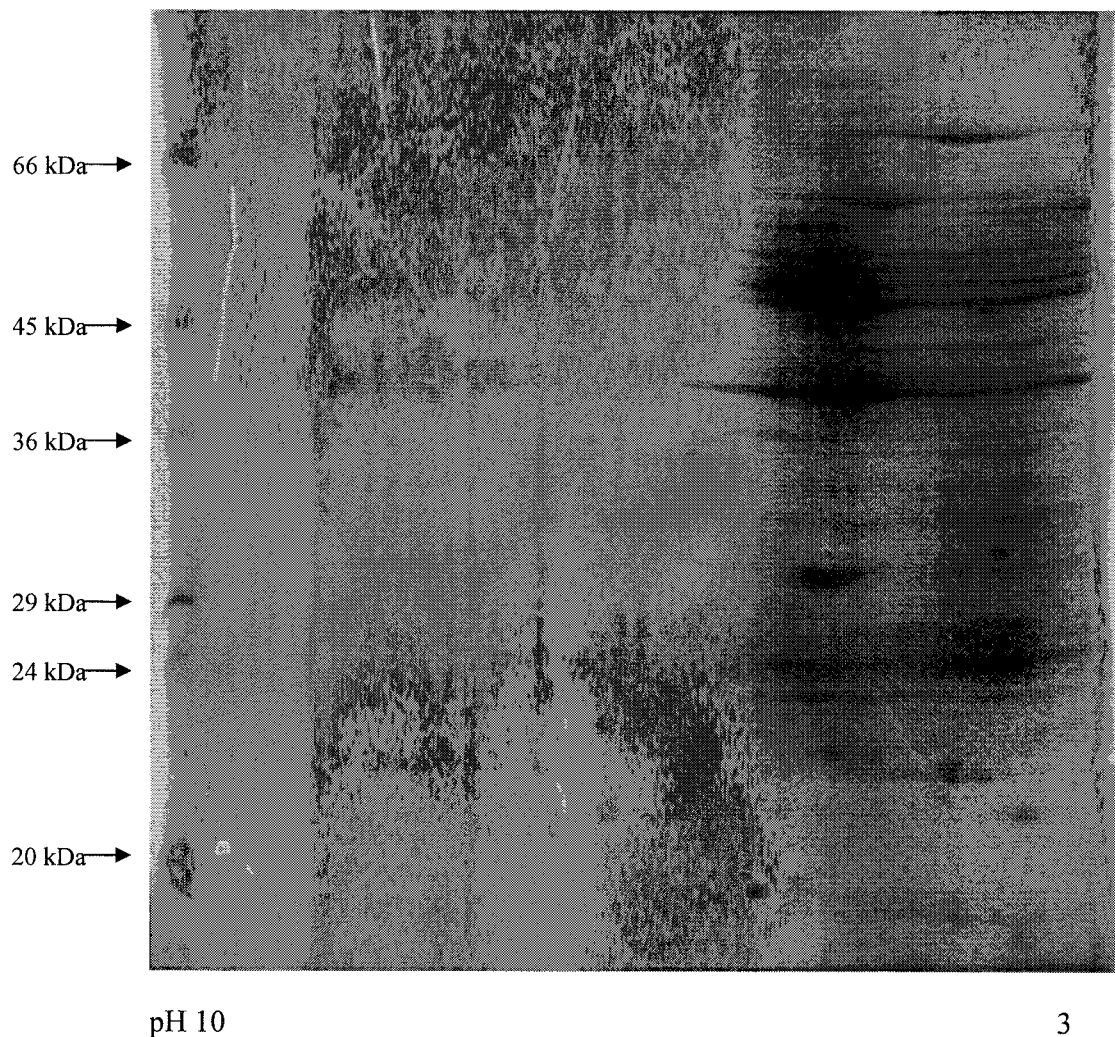


Figure 3.7: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.6).

(a) Immobiline DryStrips pH 3 -10, 18 cm were employed for IEF, and 12 % (w/v) SDS-PAGE was used for the second dimension; (b) Lysis solution B (double concentration of CHAPS) was used to extract the proteins from cell lysate; (c) Rehydration solution A was used in the rehydration step; (d) Coomassie blue stain was used to visualize the gel; (e) LMW markers: 66, 45, 36, 29, 24 and 20 kDa.

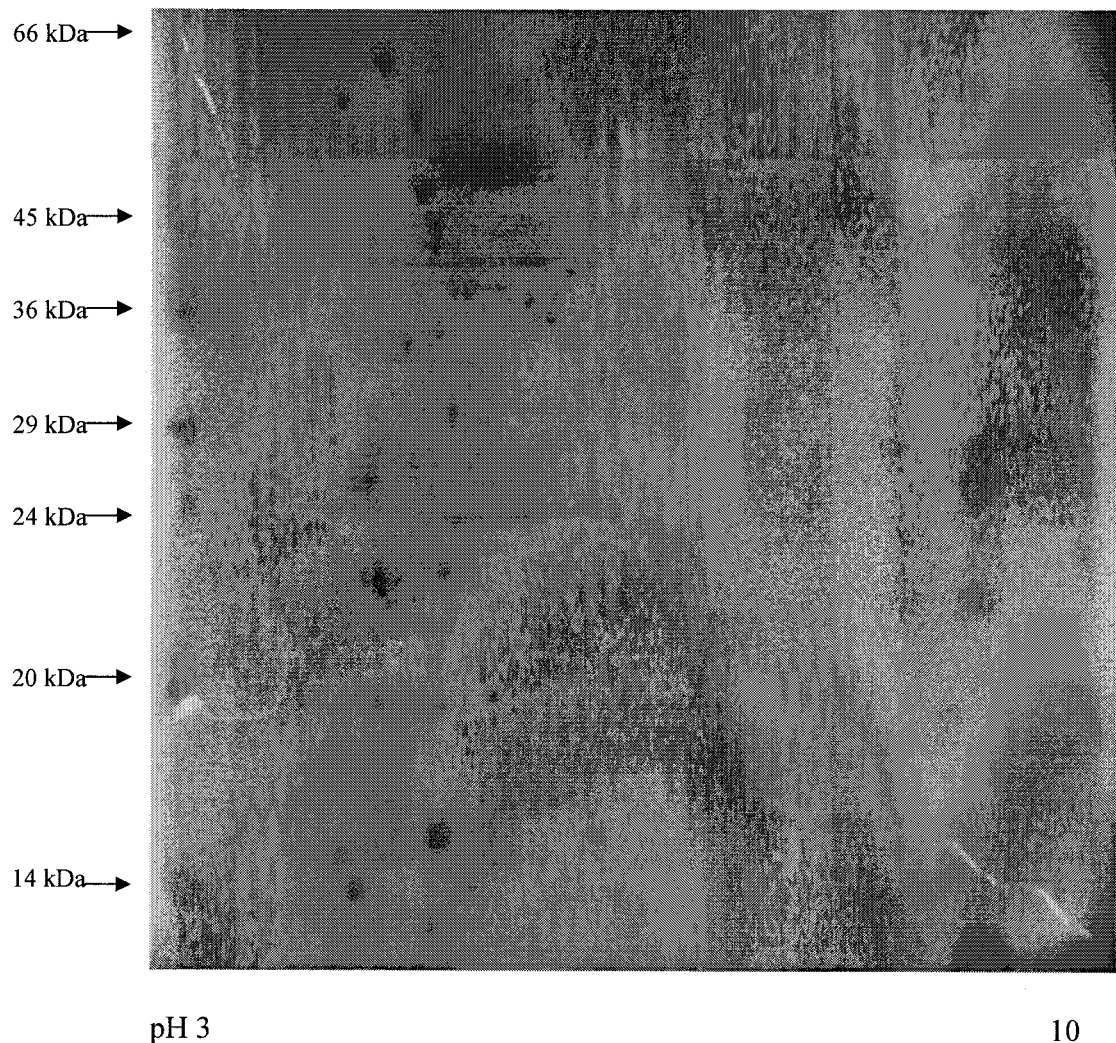


Figure 3.8: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.7).

(a) Immobiline DryStrips pH 3 -10, 18 cm were employed for IEF, and 12 % (w/v) SDS-PAGE was used for the second dimension; (b) Lysis solution B was used to extract the proteins from cell lysate; (c) a 2-D Clean-Up kit was used to remove contaminantss in the sample before loading the gel; (d) Rehydration solution A was used in the rehydration step; (e) Coomassie blue stain was used to visualize the gel; (f) LMW markers: 66, 45, 36, 29, 24, 20 and 14 kDa).

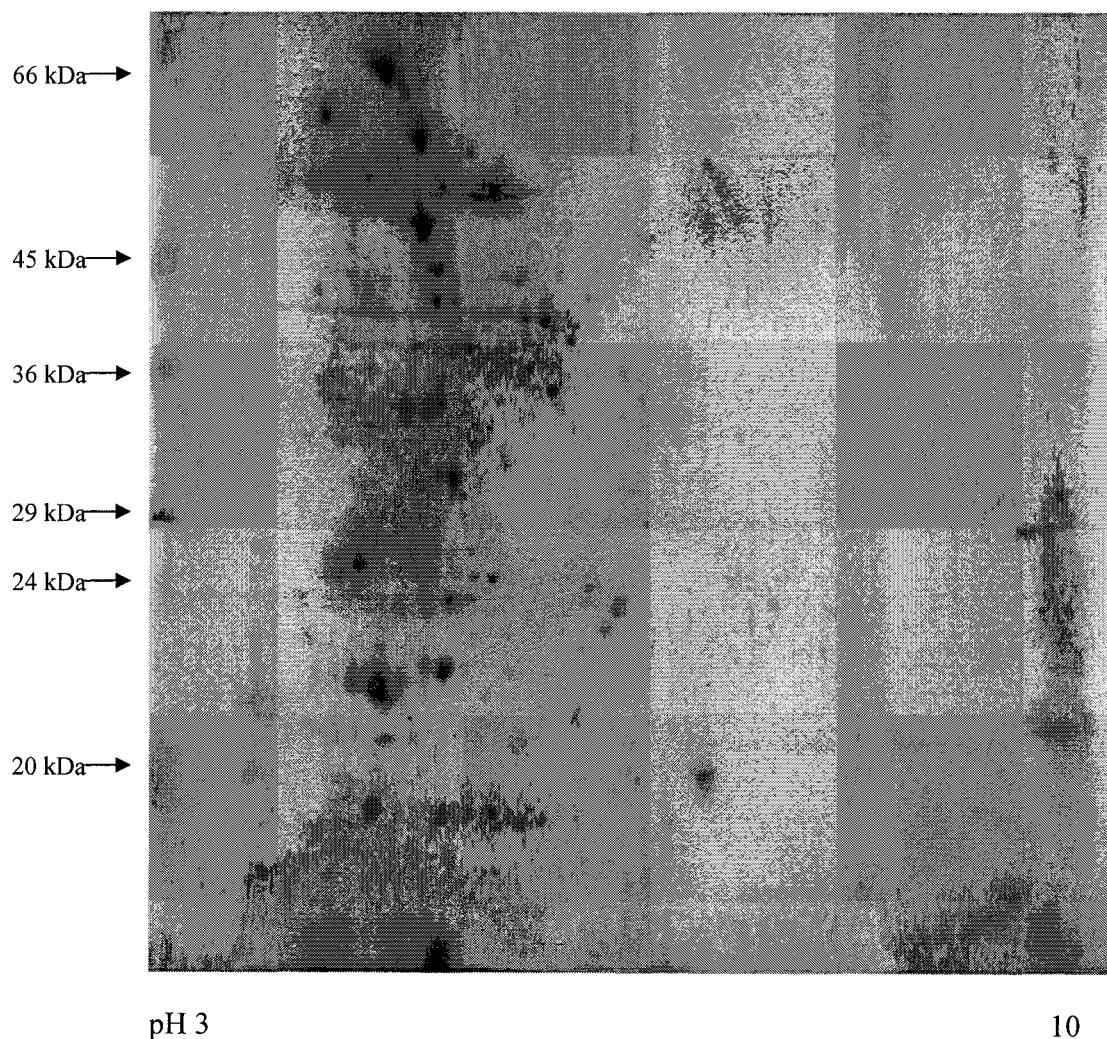


Figure 3.9: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.8).

(a) Immobiline DryStrips pH 3 -10, 18 cm were employed for IEF, and 12 % (w/v) SDS-PAGE was used for the second dimension; (b) Lysis solution B was used to extract the proteins from cell lysate; (c) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (d) Rehydration solution A was used in the rehydration step; (e) colloidal Coomassie blue stain was used to visualize the gel; (f) LMW markers: 66, 45, 36, 29, 24, and 20 kDa.

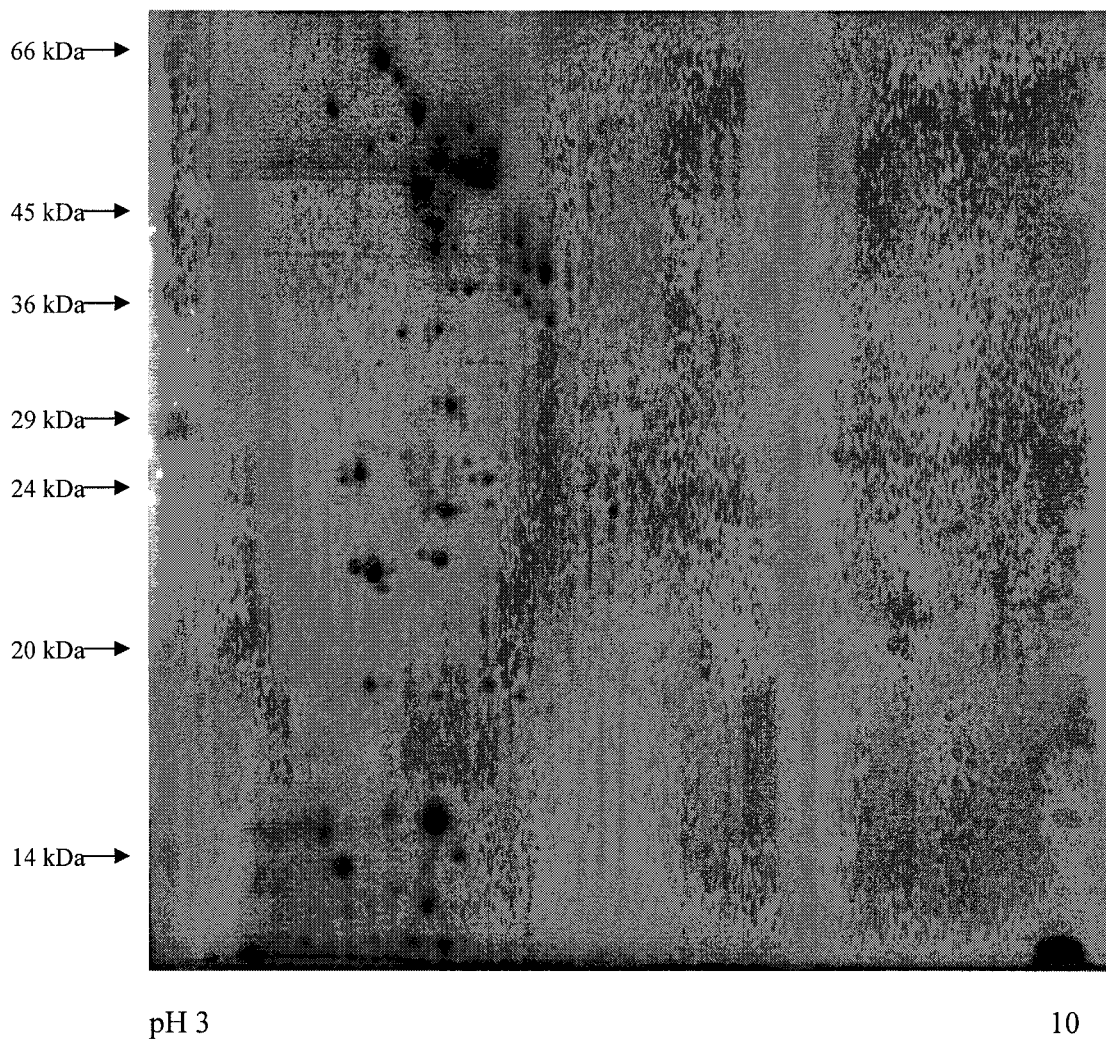


Figure 3.10: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.9).

(a) Immobiline DryStrips pH 3 -10, 18 cm were employed for IEF, and 12 % (w/v) SDS-PAGE was used for the second dimension; (b) Lysis solution B was used to extract the proteins from cell lysate; (c) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (d) Rehydration solution B (double concentration of DTT) was used in the rehydration step; (e) colloidal Coomassie blue stain was used to visualize the gel; (f) LMW markers: 66, 45, 36, 29, 24, 20 and 14 kDa.

3.2.2.1.2 Separation of cell-associated proteins in pH range 4 – 7

Analysis of the 2DE gel images, which were in pH range 3 -10, showed that greater than 90 % of cell-associated proteins from *S. pyogenes* focus in the low pH range. Therefore pH range 4-7 is a more appropriate pH range to use to analyse protein expression. Three experiments were carried out to optimize the techniques. The common conditions in these experiments were (a) Immobiline DryStrips pH 4 - 7, 18 cm were employed; (b) Lysis solution B was used to extract the proteins from cell lysate; (c) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (d) Rehydration solution B was used in the rehydration step; (e) protein samples were loaded by rehydration loading; (f) fresh Immobiline DryStrip Cover Fluid was used to cover the Immobiline DryStrips during IEF; (g) colloidal Coomassie blue staining was used to visualize the gel. The different conditions were as following:

Experiment No. 10 (Figure 3.11)*: (a) 40 ml of cell culture was used to extract the proteins; (b) 12 % (w/v) of SDS-PAGE was used for the second dimension of electrophoresis; (c) the image was acquired using the CCD camera of the Gel Doc 2000 system.

Experiment No. 11 (Figure 3.12)*: (a) 80 ml of cell culture was used to extract the proteins; (b) 12 % (w/v) of SDS-PAGE was used for the second dimension of electrophoresis; (c) the image was acquired using the CCD camera of the Gel Doc 2000 system.

Experiment No. 12 (Figure 3.13): (a) 80 ml of cell culture was used to extract the proteins; (b) 14 % (w/v) SDS-PAGE was used for the second dimension of electrophoresis; (c) the image was acquired using GS710 image densitometer.

In the experiment No.10, 40 ml of cell culture was used to extract the proteins. The 2DE gel in Figure 3.11 showed a few clear protein spots. The concentrations of DTT,

* The replicates of the 2D gels are shown in Figure E2.

urea, and CHAPS in the sample were appropriate as the image showed good protein solubilisation. Therefore the possible reason for the low number of protein spots was that not enough cells were used for sample preparation. In the experiment No. 11, 80 ml of cell culture was used to extract the proteins. The separation of proteins on the 2DE gel (Figure 3.12) is an appropriate image for analysis of protein expression. In experiment No. 12, the image was optimized by running 14 % (w/v) instead of 12 % (w/v) SDS-PAGE in the 2nd dimension and was acquired using the GS-710 image densitometer. Both changes improved the 2DE gel image and made the gels more suitable for the differential protein expression analysis.

Therefore, for analysis of differentially expressed proteins of GAS growing in BHI and BHI + HA, we used following optimized solutions and method:

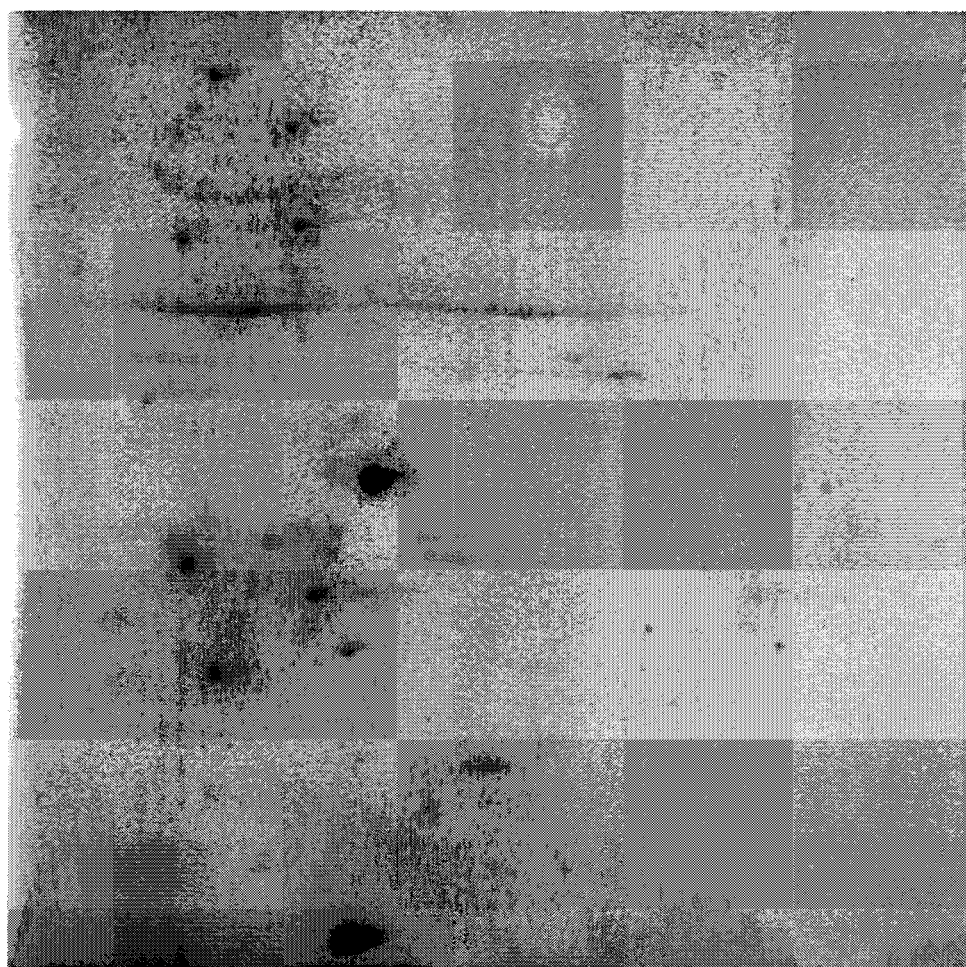
Lysis solution:

Urea	8 M
CHAPS	4 %
IPG buffer	2 %

Rehydration Solution:

Urea	8 M
CHAPS	4 %
BPB	0.002 %
IPG buffer	2 %
DTT	0.4 %

IEF in pH rang 4 – 7 was used as 1st dimension to separate cell-associated protein extract from 80 ml of cell culture. The protein extracts was treated with a 2-D Clean-Up kit, and samples were loaded via rehydration loading method. In the 2nd dimension, 14 % of SDS-PAGE was employed, and gels were visualized using colloidal Coomassie blue staining. The 2D gel images were acquired using the GS-710 image densitometer for further analysis.



pH 4

7

Figure 3.11: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.10).

(a) Immobiline DryStrips pH 4 -7, 18 cm were employed for IEF, and 12 % (w/v) SDS – PAGE was used for the second dimension; (b) 40 ml of cell culture was used to extract proteins; (c) Lysis solution B was used to extract the proteins from cell lysate; (d) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (e) Rehydration solution B was used in the rehydration step; (f) colloidal Commassie blue stain was used to visualize the gel. Size standards could not be visualised on this gel.

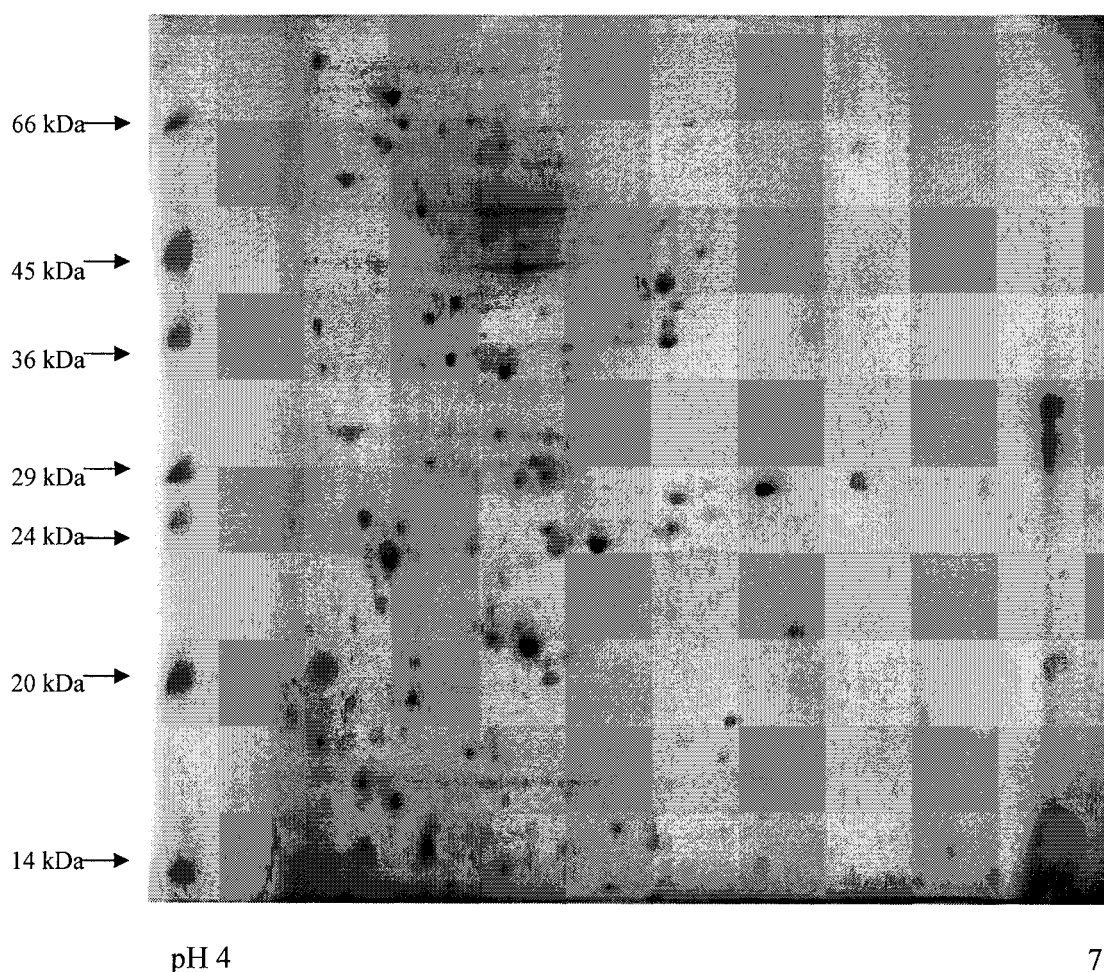


Figure 3.12: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.11)

(a) Immobiline DryStrips pH 4 -7, 18 cm were employed for IEF, and 12 % (w/v) SDS – PAGE was used for the second dimension; (b) 80 ml of cell culture was used to extract proteins; (c) Lysis solution B was used to extract the proteins from cell lysate; (d) a -D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (e) Rehydration solution B was used in the rehydration step; (f) colloidal Coomassie blue stain was used to visualize the gel; (g) LMW markers: 66, 45, 36, 29, 24, 20 and 14 kDa.

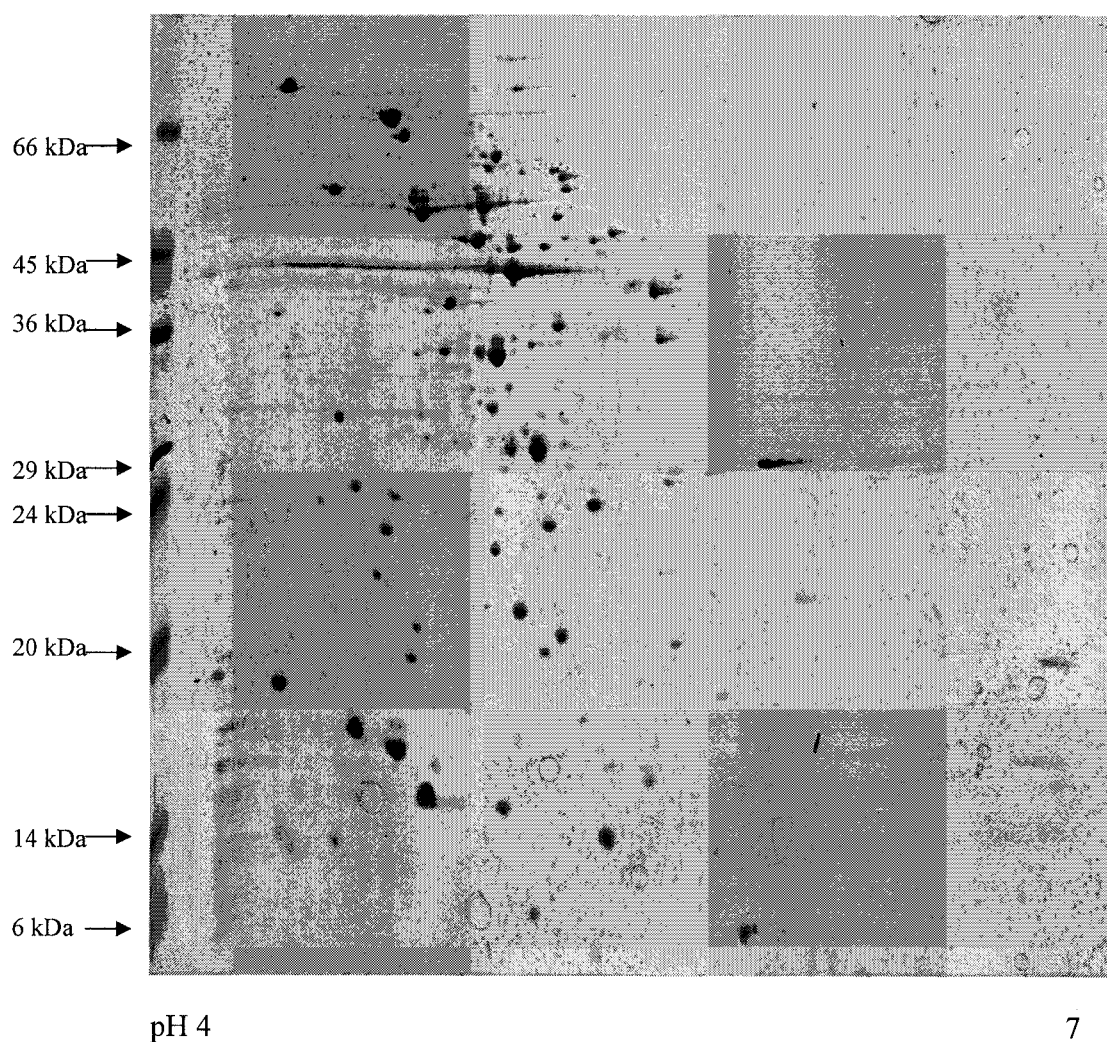


Figure 3.13: 2DE gel of cell-associated proteins from *S. pyogenes* SF370 (Experiment No.12).

(a) Immobiline DryStrips pH 4 -7, 18 cm were employed for IEF, and 14 % (w/v) SDS – PAGE was used for the second dimension; (b) 80 ml of cell culture was used to extract proteins; (c) Lysis solution B was used to extract the proteins from cell lysate; (d) 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (e) Rehydration solution B was used in the rehydration step; (f) colloidal Commassie blue stain is used to visualize the gel; (g) LMW marker: 66, 45, 36, 29, 24, 20, 14 and 6 kDa); (h) the image was acquired using the GS-710 image densitometer.

3.2.2.1.3 Separation of secreted proteins from culture supernatant in pH range 3 – 10

In experiment No.13, the secreted proteins were isolated from *S. pyogenes* SF370 culture supernatant according to the method in section 2.6.1.4, and separated using Immobiline DryStrips pH 3 - 10, 18 cm in the first dimension (Figure 3.14). The technical details are as follows: (a) 25 ml of culture supernatant was used to obtain proteins; (b) a 2-D Clean-Up kit was used to remove contaminants in the sample before loading the gel; (c) Rehydration solution B was used in the rehydration step; (d) protein samples were loaded by cup loading rehydration loading; (e) the fresh Immobiline DryStrip Cover Fluid is used to cover the Immobiline DryStrips during IEF; (f) 12 % (w/v) SDS-PAGE was used in the second dimension of electrophoresis; (g) colloidal Coomassie blue staining was used to visualize the gel; (h) the image was acquired using the GS-710 image densitometer.

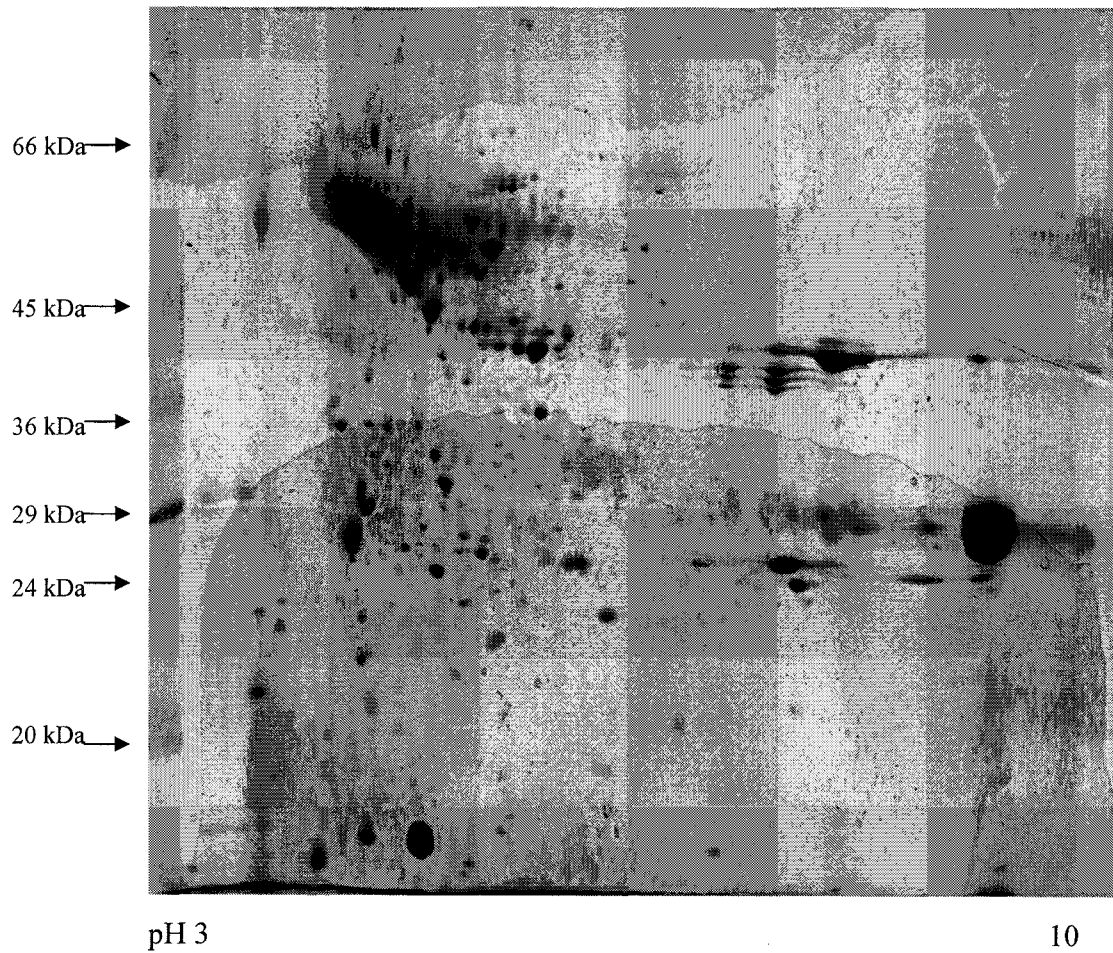


Figure 3.14: 2DE gel of secreted proteins from *S. pyogenes* SF370 culture supernatant (Experiment No.13).

(a) Immobiline DryStrips pH 3 - 10, 18 cm were employed for IEF, and 12 % (w/v) SDS – PAGE was used for the second dimension; (b) 25 ml of cell culture supernatant was used to obtain proteins; (c) LMW markers: 66, 45, 36, 29, 24, and 20kDa.

3.3 Development of static nanospray MS analysis and protein identification using BioWorks v 3.2 software

3.3.1 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS

To analyse the salt and detergent tolerance of the static nanospray LCQ Advantage, nine proteins (β -Galactosidase, phosphorylase b, albumin, L-glutamic dehydrogenase, ovalbumin, glyceraldehyde-3-phosphate dehydrogenase, carbonic anhydrase, trypsinogen and α -lactalbumin) were trypsin digested (Figure 3.15). The digests were desalted using ZipTips and eluted with solution containing various concentration of NaCl or SDS (section 2.9.3).

10 μ g of each sample was analysed using static nanospray LCQ Advantage MS, and the spectrum were collected for 100 scans. Figure 3.16 shows the MS profiles for the albumin digestion containing varying concentration of salt or SDS. The MS profiles for the remaining proteins are shown in the Appendix D

3.3.2 Purified protein identification using static nanospray LCQ Advantage MS with BioWorks 3.2 software

The purified protein, unsaturated glucuronyl hydrolase (Ugl) from *S. pyogenes* SF370 was digested according to the method in section 2.8.1 (Figure 3.17), and the digestion was analysed using static nanospray LCQ Advantage MS (Figure 3.18). The spectrum was loaded on BioWorks 3.2 software. The MS/MS data was searched against the streptococci database (section 2.10.1.1) using TurboSEQUEST search programme. Ugl was present as the first hit in the result table and the data was of high quality (Table 3.1). Table 3.2 shows the coverage of this identified protein.

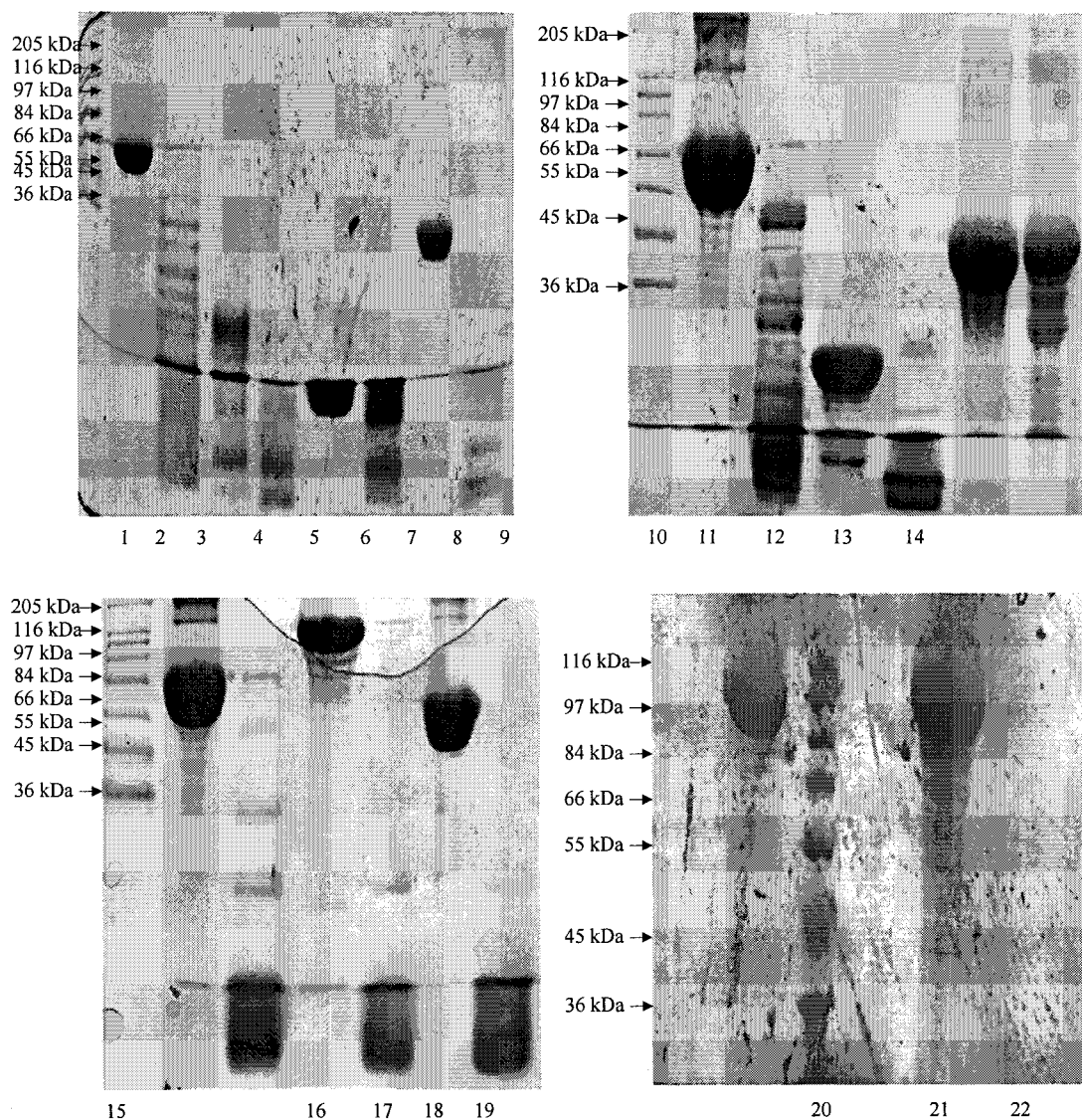
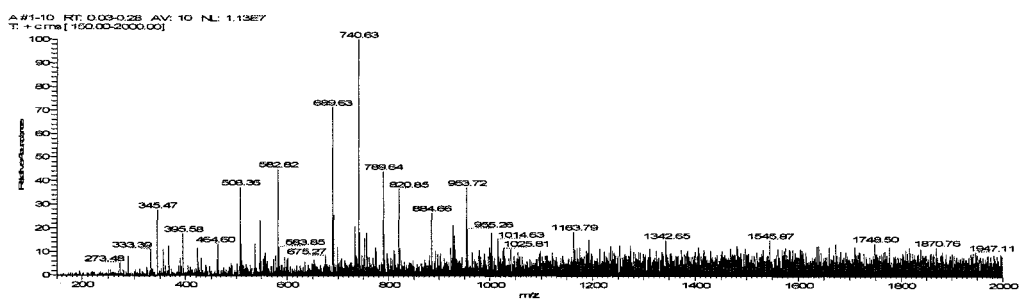


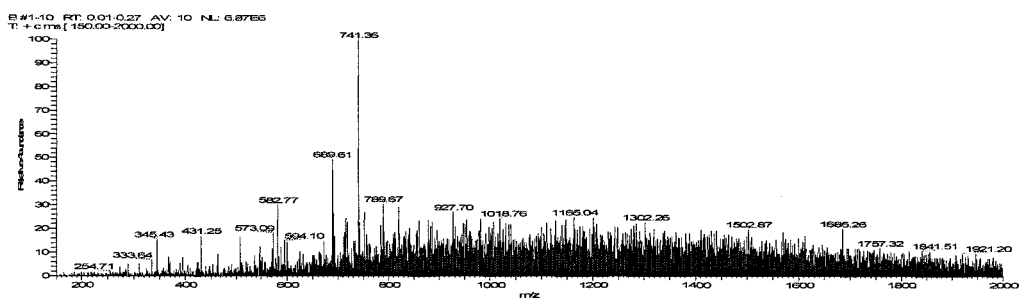
Figure 3.15 12 % (w/v) SDS-PAGE analysis of tryptic digestion of proteins

Lane 1, 10, 15 and 20: high molecular weight (HMW) makers: 205, 116, 97, 84, 66, 55, 45 and 36 kDa. Lane 2: L-glutamic dehydrogenase protein; Lane 3: L-glutamic dehydrogenase digestion; Lane 4: trypsinogen protein; Lane 5: trypsinogen digestion; Lane 6: α -lactalbumin protein; Lane 7: α -lactalbumin digestion; Lane 8: glyceraldehyde-3-phosphate dehydrogenase; Lane 9: Glyceraldehyde-3-phosphate dehydrogenase digestion; Lane 11: albumin; Lane 12: albumin digestion; Lane 13: carbonic anhydrase; Lane 14: Carbonic anhydrase digestion; Lane 16: phosphorylase b; Lane 17: phosphorylase b digestion; Lane 18: ovalbumin; Lane 19: ovalbumin digestion; Lane 21: β -galactosidase ; Lane 22: β -galactosidase digestion.

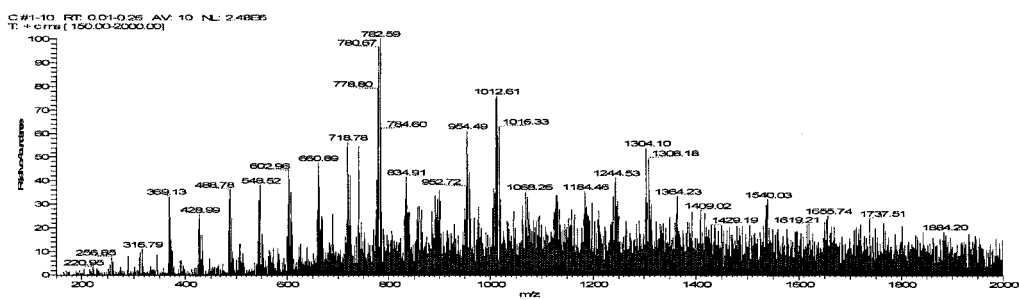
A



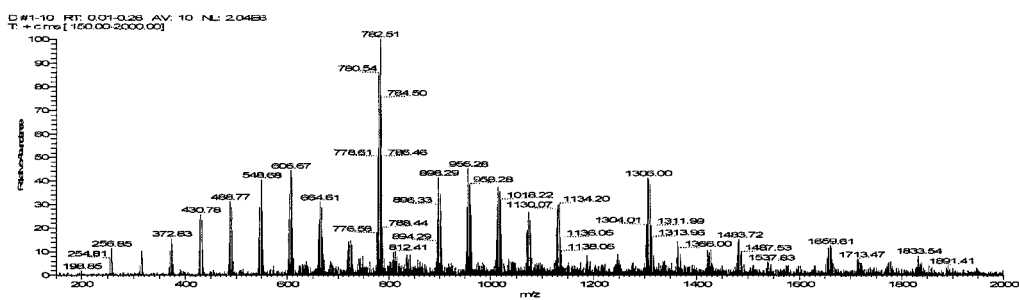
B



C

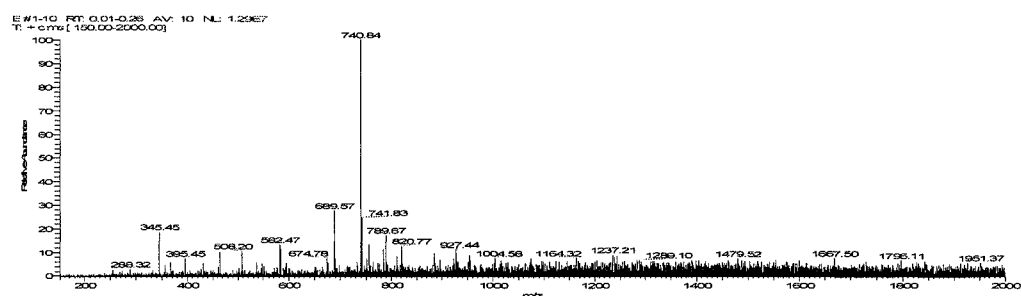


D

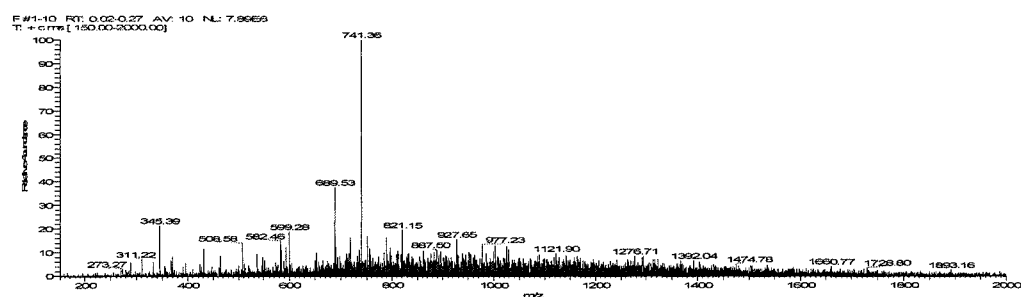


(Figure continues on next page)

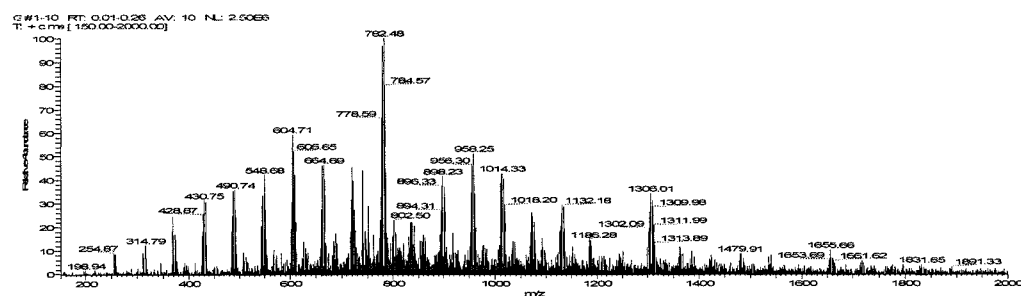
E



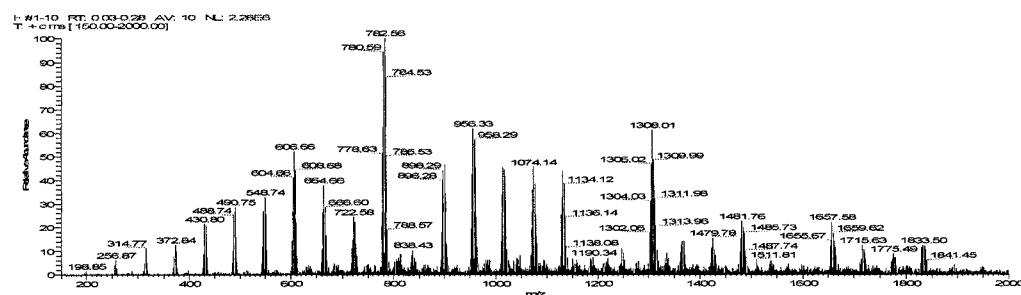
F



G

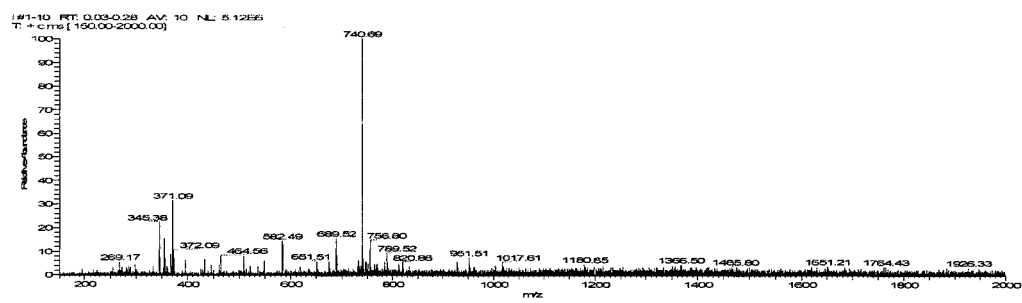


H

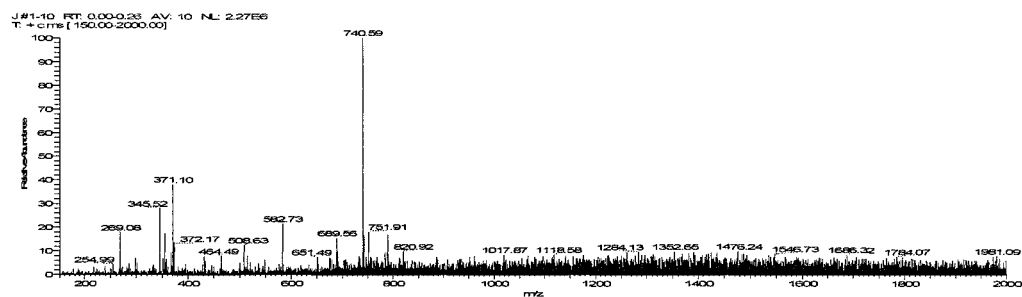


(Figure continues on next page)

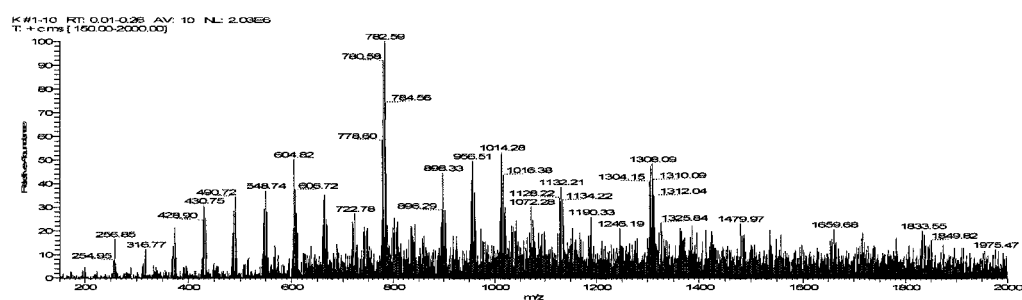
I



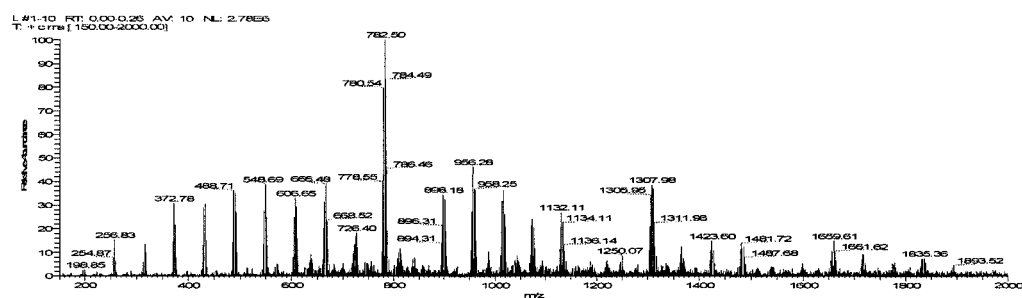
J



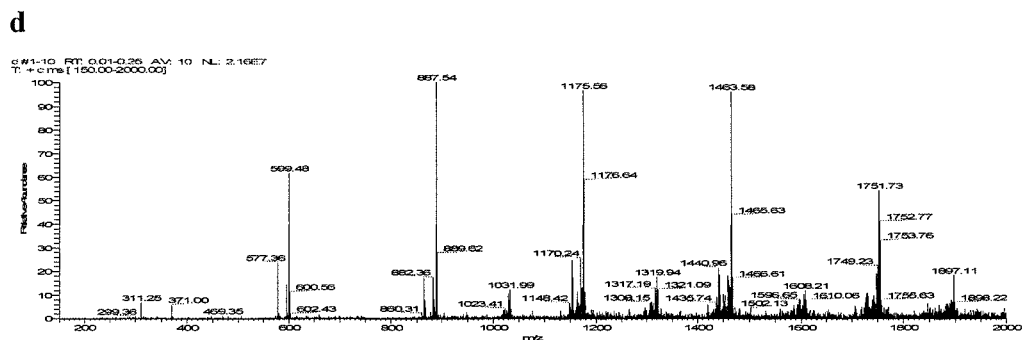
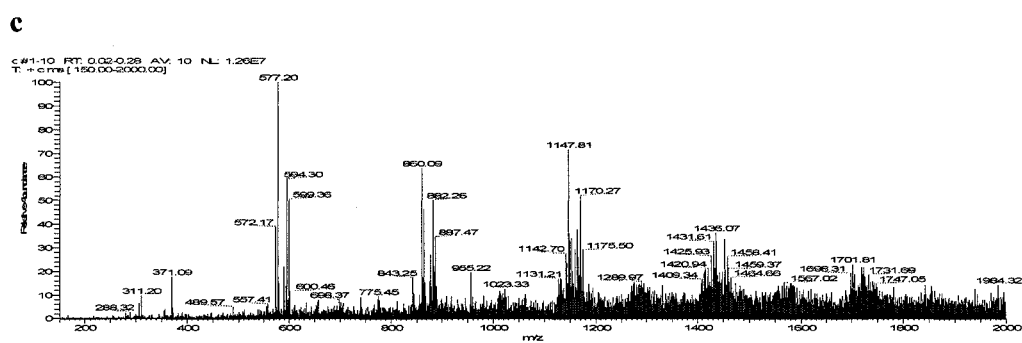
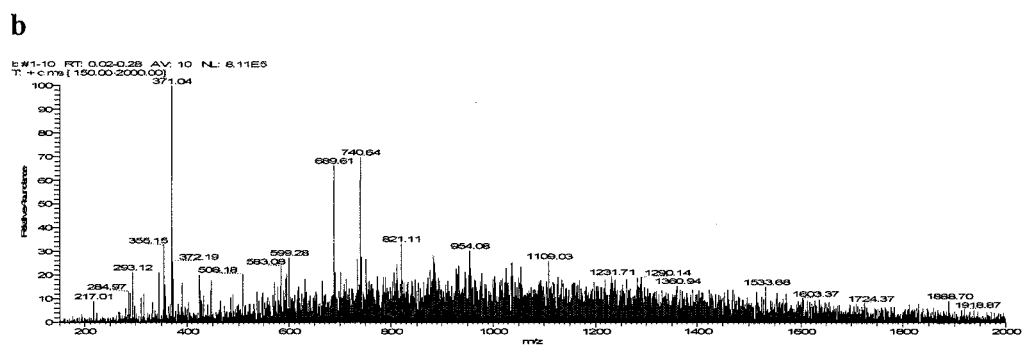
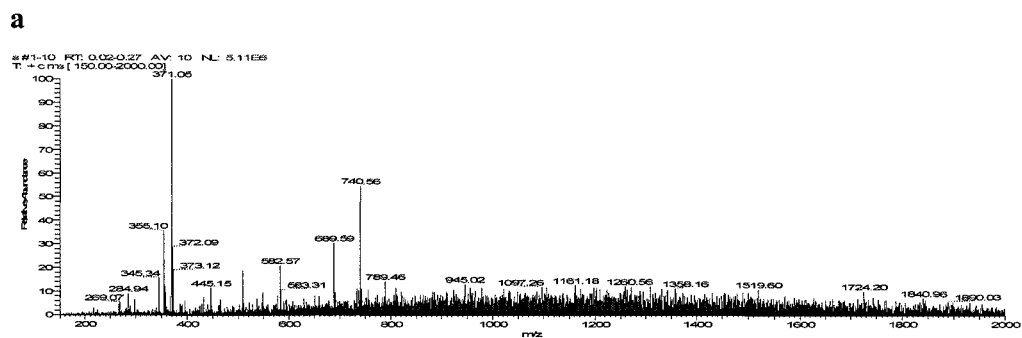
K



L

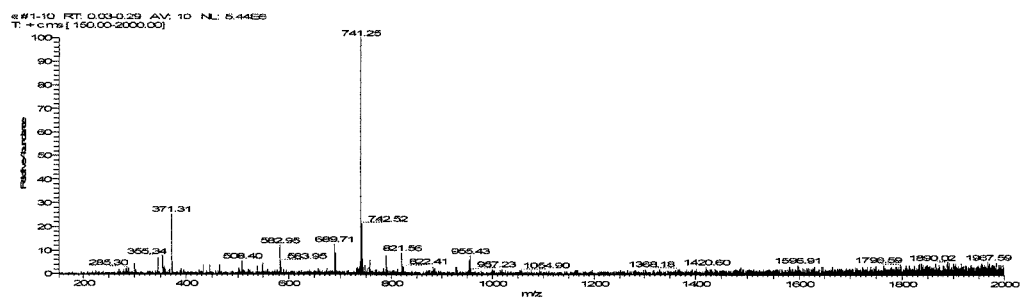


(Figure continues on next page)

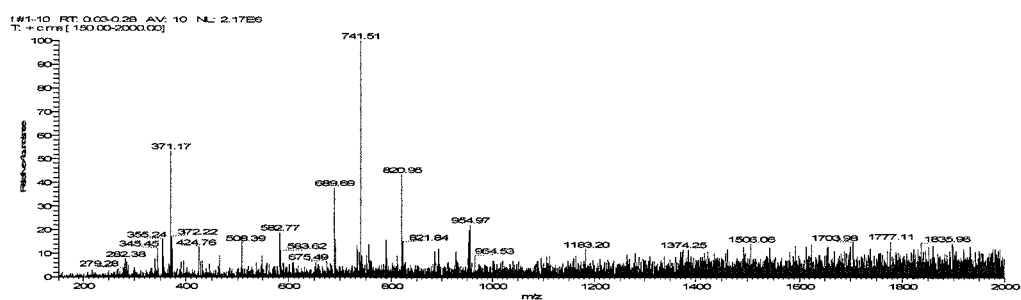


(Figure continues on next page)

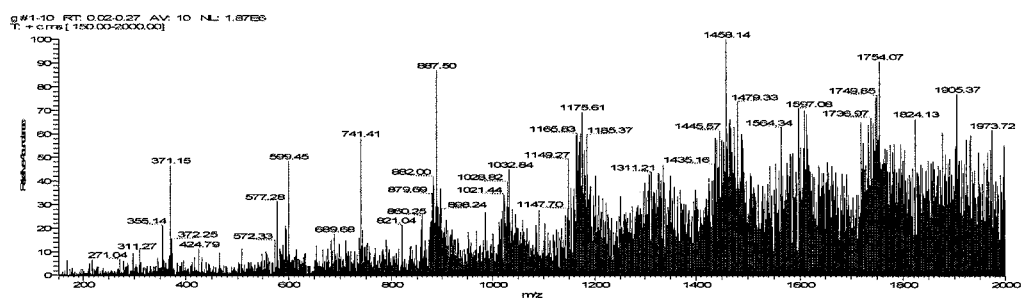
e



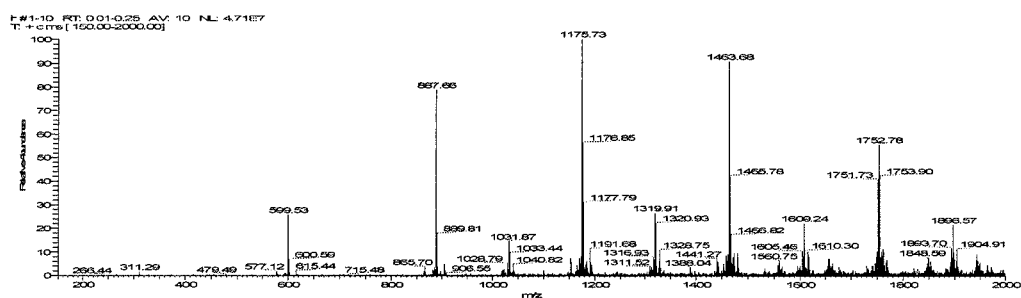
f



g



h



(Figure continues on next page)

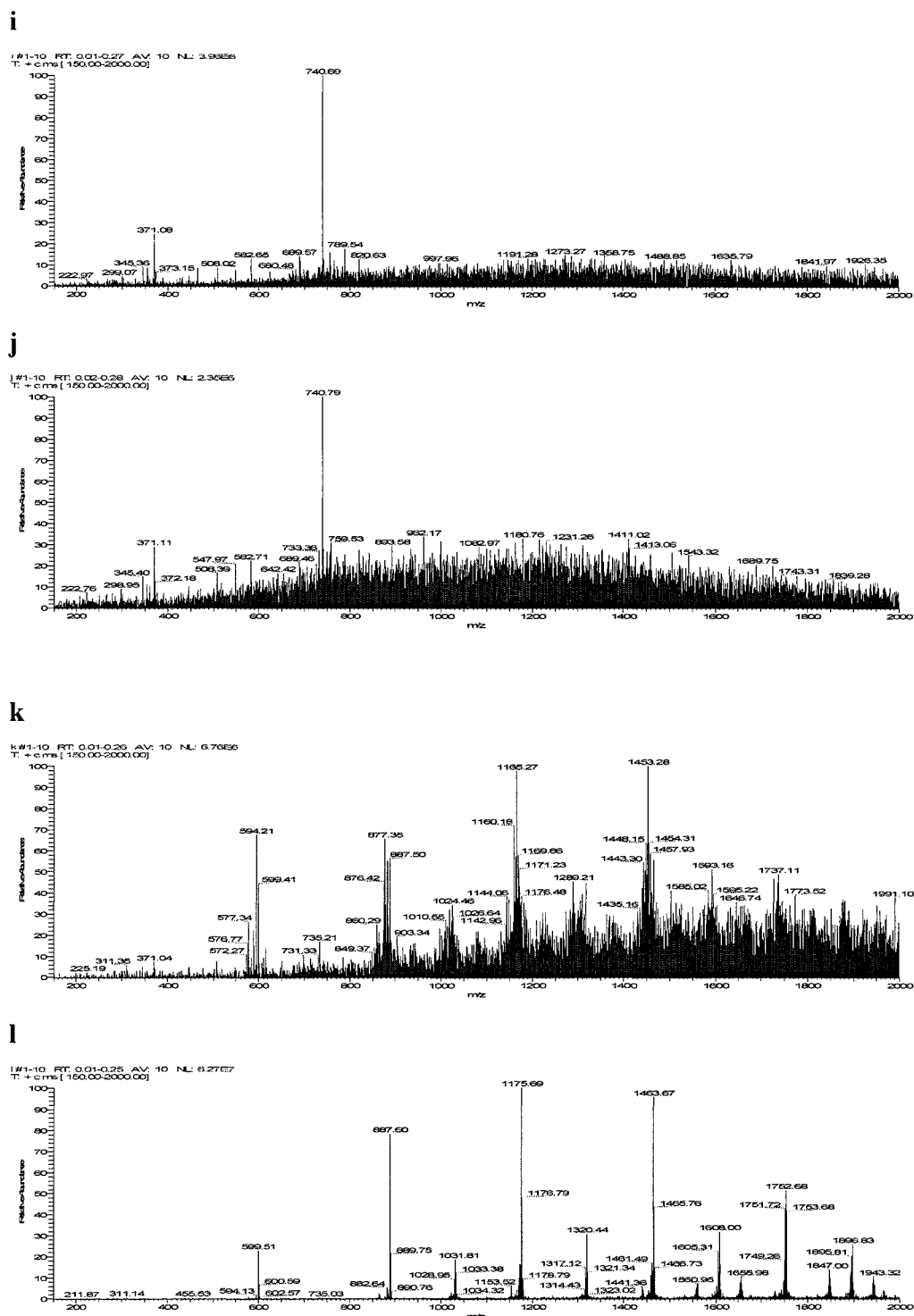


Figure 3.16 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using albumin digests.
The sample conditions (A-L and a-l) are given in Table 2.10.

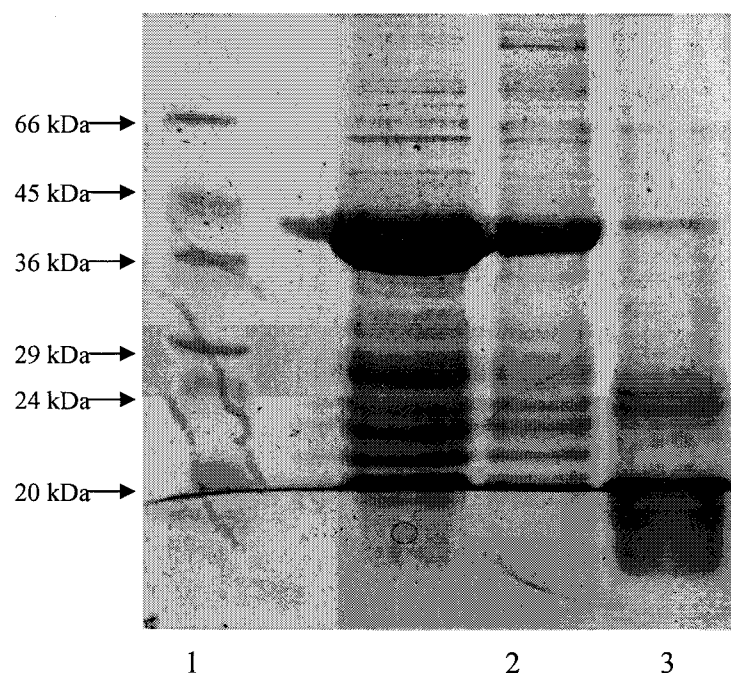


Figure 3.17 12 % (w/v) SDS – PAGE analysis of Ugl digestion.

Lane 1: LMW markers: 66, 45, 36, 29, 24 and 20 kDa; Lane 2: Ugl protein; Lane 3 Ugl digestion.

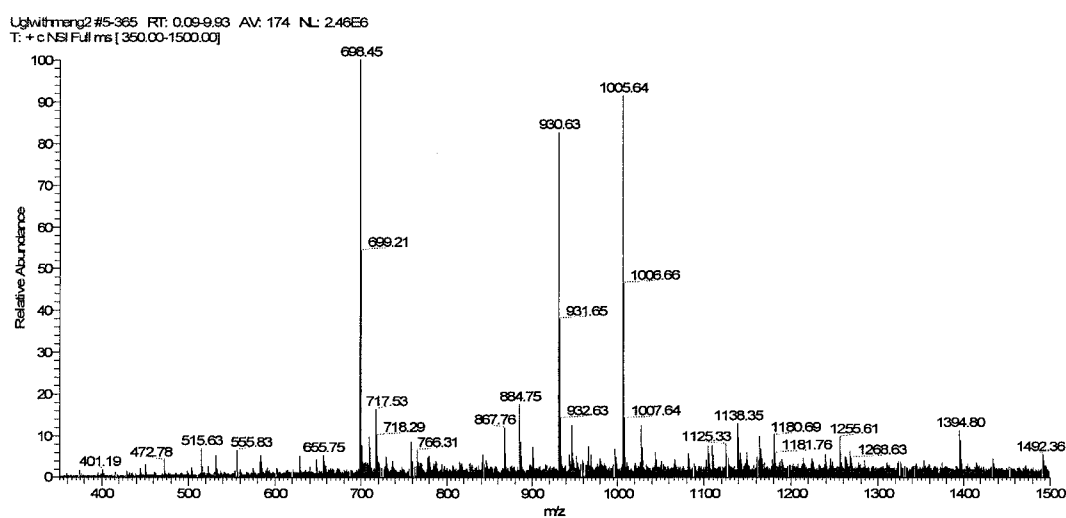


Figure 3.18 Mass spectrum of Ugl digestion

10 μ l of Ugl digestion was analysed by static nanospray LCQ Advantage MS, and the data was collected through Xcalibur software.

Scan(s)	Peptide	z	Xcorr	ΔCn
(NC_002737) putative unsaturated glucuronyl hydrolase [<i>Streptococcus pyogenes</i> M1]				
			80.28	
122 - 140	K.EDFPTPATK.D	1	1.59	0.30
142 - 144	R.QGYSDDSAWAR.G	2	3.40	0.61
148	R.QGYSDDSAWAR.G	2	2.94	0.49
154 - 164	R.QGYSDDSAWAR.G	2	3.68	0.52
166 - 182	K.TIALEPIK.Q	1	2.44	0.29
184	K.TIALEPIKQPER.F	2	3.85	0.47
186	K.TYEAAMHSMLR.A	2	3.76	0.60
188 - 214	K.TIALEPIKQPER.F	2	3.92	0.51
216	R.LNMDYFK.E	1	2.08	0.13
218	K.TIALEPIKQPER.F	2	3.46	0.44
220 - 234	R.LNMDYFK.E	1	1.98	0.15
236 - 238	K.TIALEPIKQPER.F	2	4.07	0.56
240 - 246	K.TIALEPIK.Q	1	2.30	0.29
248	K.TIALEPIKQPER.F	2	3.71	0.55
252	K.TIALEPIKQPER.F	2	4.01	0.51
254 - 276	R.LNMDYFK.E	1	1.88	0.13
278 - 308	K.TIALEPIKQPER.F	2	4.09	0.54
310 - 314	K.ALAQANDLSFLDR.V	2	5.59	0.59
316	K.TIALEPIKQPER.F	2	4.32	0.55
318	K.ALAQANDLSFLDR.V	2	5.27	0.65
320 - 330	K.TIALEPIKQPER.F	2	3.68	0.50
332 - 334	K.TYEAAMHSMLR.A	2	4.11	0.64
342 - 344	R.DMAINHFYASANHVR.D	3	5.30	0.65

Table 3.1 Information of matched peptides from Ugl

Sequence	MH+	% by Mass	Position	% by AA
TIALEPIK	884.54515	1.92	7 - 14	2.01
TIALEPIKQPER	1394.8002	3.03	7 - 18	3.01
LNMDYFK	930.43897	2.02	43 - 49	1.76
EDFPTPATK	1005.4888	2.19	50 - 58	2.26
ALAQANDLSFLDR	1433.7383	3.12	93 - 105	3.26
DMAINHFYASANHVR	1858.9017	4.04	198 - 213	4.02
QGYSDDSAWAR	1255.5338	2.73	238 - 248	2.76
TYEAAMHSMLR	1309.6028	2.85	332 - 342	2.76
Total		21.9		21.84

Table 3.2 Protein coverage of the matched peptides from Ugl

4 Results for proteomic analysis of streptococci grown in HA-rich media

4.1 Background

S. pyogenes is human pathogen. To investigate the expression of virulence factors, the bacteria were grown in media containing HA, which is a major component of mammalian connective tissue. The growth of bacteria in defined media and non-defined media was investigated. The effect of the presence of HA in non-defined media on bacteria growth was determined, including the expression of hyaluronate lyases which play an important role in degrading host tissue HA and the consumption of carbohydrates as carbon/energy sources. The comparison of differentially expressed proteins of streptococci was carried out by running 2DE gels to separate cell-associated proteins extracted from cultures grown in two different media (i.e. with and without added HA) using identical experimental conditions, including loading the same amount of the protein on each gel. The differentially expressed proteins from *S. pyogenes* SF370, which were determined using PDQuestTM v 6.2.1 software according to Wilcoxon two-sample test, were analysed and identified using static nanospray LCQ Advantage MS with BioWorks 3.2 software. In order to identify and understand the functions of differentially expressed proteins that may be relevant to GAS pathogenesis, TSSViewTM v1.0 was used in this study. Two proteins that could not be identified using TurboSEQUENT according to the predefined criteria, i.e. proteins must be from the SF370 strain of the M1 serotype of GAS; and peptides must have a Xcorr > 1.5 for singly charged peptides, > 2.0 for doublely charged peptides, and > 2.5 for triplely charged peptides, with at least 2 peptides successfully matched; and/or must have a $\Delta C_n > 0.2$ with at least 2 peptides successfully matched, were added to the list of identified proteins as TSSViewTM software showed that the single high quality peptides identified from these proteins were from highly-conserved protein sequences. In addition, the potential functions of hypothetical proteins were investigated using the TSSviewTM software. Additional potential functions of identified proteins that may be relevant to pathogenesis were assigned by identifying highly-conserved region in the proteins using TSSviewTM software.

4.2 Proteomic analysis of GAS grown in HA-rich media

4.2.1 Growth of streptococci in defined media

The defined media used in this project has been shown to be a good media as several oral streptococci have been shown to grow as rapidly as they do in a non-defined media such as Todd- Hewitt broth when incubated either aerobically or anaerobically (Homer et al., 1993, Terleckyj et al., 1975). It was chosen as a basal media, to which either glucose or HA was added as the source of carbon/energy (section 2.1.4 and section 2.2.5.1). Glucose or HA (0.02 (w/v) %, 0.1 (w/v) % and 0.2 (w/v) %) was added to the basal media and the growth was monitored every 24 hours by measuring the OD at 660 nm against corresponding media (Table 4.1).

As *S. pyogenes* SF370 showed little growth in defined media with various concentration of glucose, and no growth in defined media with various concentration of HA the effect of the presence of HA in the media on the differential protein expression of streptococci could not be carried out using these media. Therefore, a non-defined medium, BHI, was employed for proteomic studies.

4.2.2 Growth of *S. pyogenes* SF370 in HA-rich media and control media

BHI as a non-defined media that has been employed in many streptococci studies (Homer et al., 1994, Wilkins et al., 2002). *S. pyogenes* SF370 showed rapid growth in BHI broth (control media) and BHI broth with 0.5 % (w/v) HA (HA-rich media) (section 2.1.5) under aerobic condition, and the growth rate was monitored (Figure 4.1)

	0.02 %	0.1 %	0.2 %	0.02 %	0.1 %	0.2 %
	(w/v)	(w/v)	(w/v)	(w/v)	(w/v)	(w/v)
	Glucose	Glucose	Glucose	HA	HA	HA
24 h	0.179	0.183	0.192	0.014	0.009	0.010
48 h	0.182	0.187	0.184	0.012	0.013	0.015
72 h	0.171	0.180	0.189	0.018	0.011	0.013
96 h	0.174	0.179	0.180	0.013	0.012	0.009

Table 4.1 Growth of *S. pyogenes* SF370 in defined media

The OD at 660nm of *S. pyogenes* SF370 in six defined media were measured at 24, 48, 72 and 96 h.

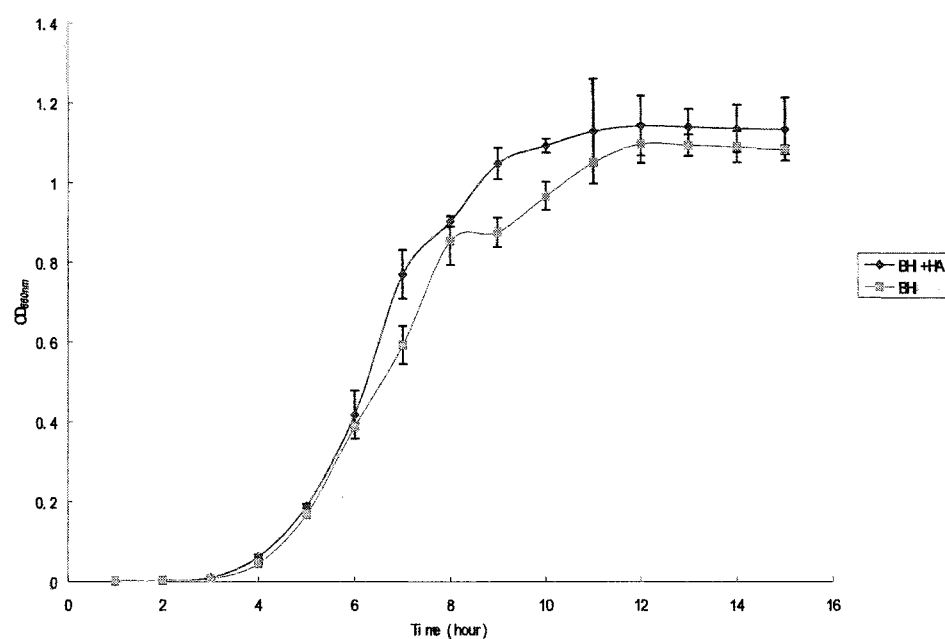


Figure 4.1 Growth curves for *S. pyogenes* SF370 in HA-rich media (♦) and BHI media (■).

The growth curves were monitored by measuring OD at 660nm. All data were performed in triplicate with error bars representing standard deviation from the mean.

4.2.3 Investigation into the expression of hyaluronate lyases in *S. pyogenes* SF370 cultures

The degradation of HA in the host connective tissue by virulent streptococci is a key element of invasive infection (Hynes and Ferretti, 1989). Pathogens produce hyaluronate lyases to depolymerise host connective tissue HA (Li et al., 2000). To investigate if *S. pyogenes* SF370 produced any hyaluronate lyases due to growth in the presence of HA in BHI media, DNSA reducing sugar and 232 nm assays were employed using culture supernatant as the ‘enzyme’. The reactions for DNSA assay were carried out at optimum conditions for HyIP1, a *S. pyogenes* hyaluronate lyase (Smith, 2004) against dialysed substrate sodium hyaluronate. The release of reducing sugar was measured at 0 time point and overnight after adding ‘enzyme’ (Table 4.2).

The reducing sugar in sample nos. 1-4 was measured at 0 time point, after adding ‘enzymes’ (section 2.3.1), and the reducing sugar in sample nos. 5-8 was measured after overnight incubation. The difference between no. 1 and 5, no. 2 and 6, no. 3 and 7, and no.4 and 8 would be due to the degradation of HA for each added ‘enzyme’. The DNSA assay showed that there was no obvious difference in the amount of reducing sugar in either the HA-rich culture or the control media. Therefore there was no evidence showing that any hyaluronate lyases are produced by *S. pyogenes* SF370 due to growth in the presence of HA in the culture media.

The 232 nm assay was carried out in the same conditions as the DNSA assay, the absorbance was measured before addition of each ‘enzyme’ and after addition of each ‘enzyme’. Figure 4.2 (a and b) shows that there was no evidence of any β -elimination reactions catalyzed by hyaluronate lyases occurring after adding culture supernatants, i.e. ‘enzyme’. Thus there is no evidence that any hyaluronate lyases are produced by *S. pyogenes* SF370 due to growth in the presence of HA in the culture, which is consistent with the results from the DNSA assay.

Sample No.	Concentration of reducing sugar (mg/ml)
1	0.578 ± 0.003
2	0.530 ± 0.002
3	0.430 ± 0.006
4	0.400 ± 0.003
5	0.593 ± 0.004
6	0.544 ± 0.007
7	0.444 ± 0.002
8	0.415 ± 0.006

Table 4.2 The concentrations of reducing sugar in different samples (Table 2.4) for investigating the expression of hyaluronate lyases.

4.2.4 Investigation into the consumption of carbohydrates by *S. pyogenes* SF370

During the growth of *S. pyogenes*, carbohydrates serve as the most facile source of carbon/energy (Iyer et al., 2005). To investigate the consumption of carbohydrates in *S. pyogenes* SF370 cultures, DNSA reducing sugar assay and uronic acid assay were carried out. The changes in total reducing sugar before and after growth of *S. pyogenes* in control media and HA-rich media were determined using a DNSA assay. In the control media, the concentration of reducing sugar decreased from 93 ± 2.6 mg/ml to 35 ± 1.7 mg/ml before and after *S. pyogenes* growth, respectively; in HA-rich media, the reducing sugar decreased from 95 ± 2.1 mg/ml to 36 ± 1.9 mg/ml before and after *S. pyogenes* growth, respectively. Therefore there is no evidence to show that *S. pyogenes* SF370 consumes more carbohydrates in HA-rich media than in control media.

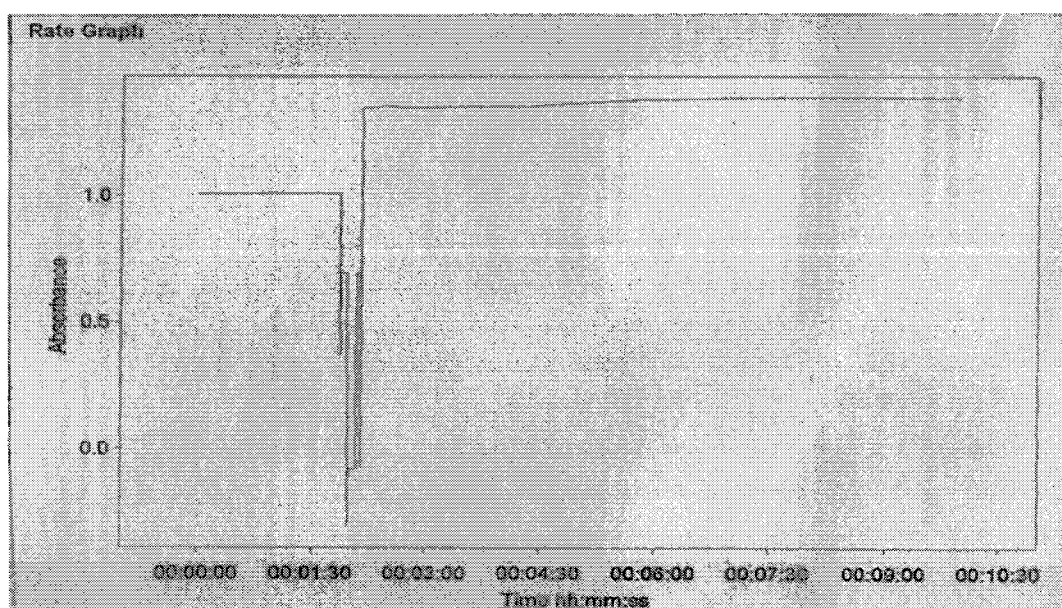
Uronic acid assay was carried out to indirectly investigate the consumption of HA, rather than carbohydrate, in *S. pyogenes* SF370 cultures. After the overnight growth, the concentrations of uronic acid decreased from 58 ± 1.2 mg/ml to 54 ± 1.5 mg/ml in control media and from 63 ± 1.0 mg/ml to 55 ± 1.3 mg/ml in HA-rich media. Therefore *S. pyogenes* SF370 consumed 4 mg/ml more uronic acid when grown in HA-rich media than in control media, this may be due to the utilization of the HA which was added to the media.

4.2.5 Comparison of differentially expressed cell-associated proteins of *S. pyogenes* SF370 growing in HA-rich and control media using 2DE*

The cell-associated proteins were extracted from *S. pyogenes* SF370 growing in either HA-rich or control media according to the method optimized in previous studies (section 3.2.2.1.2). Proteins expressed by this bacteria growing in both media were separated using 2DE, in the pH range 4 -7 using 14 % (w/v) SDS-PAGE. The optimal conditions from previous results in section 3.2.2.1.2 were applied in this section.

* The work described in this section was submitted to Proteomics and is In Press, a copy of the submitted paper can be found at the end of this thesis.

(a)



(b)

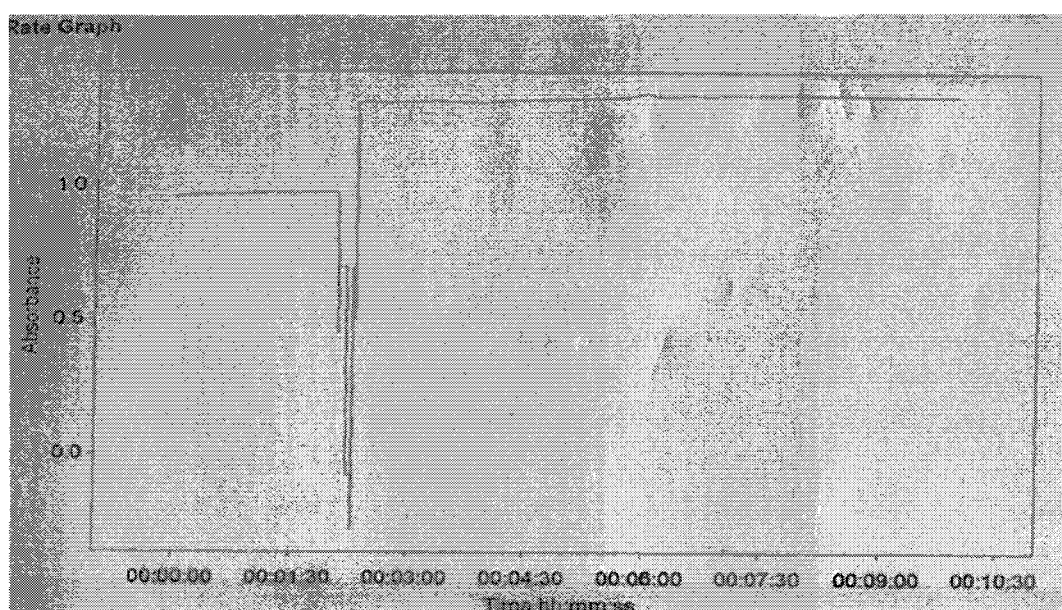


Figure 4.2 232 nm assays for investigating the expression of hyaluronate lyases.

(a) Change in absorbance after adding culture supernatants from HA-rich media. (b) Change in absorbance after adding culture supernatants from control media.

Eight pairs of the gels (i.e. control media and HA-rich media) were analysed (Houtman et al., 2003), and in each pair, a similar amount of protein, $150 \pm 10 \mu\text{g}$, was loaded on to each gel (Figure 4.3). The proteins amounts were determined by either Bradford assay or 2-D Quant Kit.

4.2.6 Analysis of protein expression of GAS growing in HA-rich media and control media using PDQuestTM v 6.2.1 software

The eight pairs of 2DE gels, for separation of cell-associated proteins from *S. pyogenes* SF370 growing in HA-rich media and control media, were analysed using PDQuestTM v 6.2.1 software. A total of 210 ± 20 and 179 ± 20 spots of cell-associated proteins from bacteria cultured in HA-rich media and control media, respectively, were detected on the 2DE gels. About 80 % of the protein spots were matched between the two sets of gels and the correlation coefficient between the two patterns was 0.815. The 18 protein spots highlighted in Figure 4.4 represent those that were up or down regulated according to the Wilcoxon two-sample test (Houtman et al., 2003). This was achieved using manual view intensities bars for each protein spot on the eight pairs of 2DE gels. The proteins where all the spots of that protein on the gels prepared from cells grown in HA-rich media were higher than all the spots of that protein on the gels prepared from cells grown in control media were defined as up-regulated proteins. Contrarily, the proteins where all the spots of that protein on the gels prepared from cells grown in HA-rich media were lower than all the spots of that protein on gels prepared from cells grown in control media were defined as down-regulated proteins. Among the 18 proteins differentially expressed, 12 proteins were up-regulated on the 2DE gels prepared from cells grown in HA-rich media and 6 proteins were down-regulated on the 2DE gels prepared from cells grown in HA-rich media.



pH 4

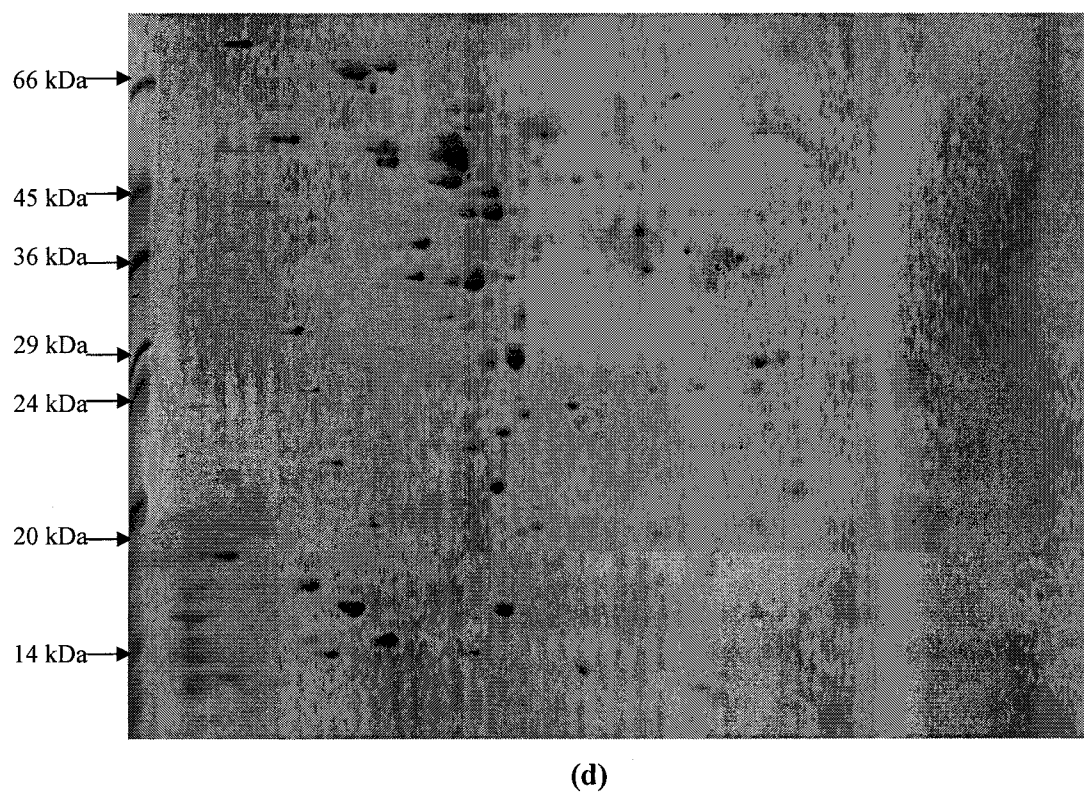
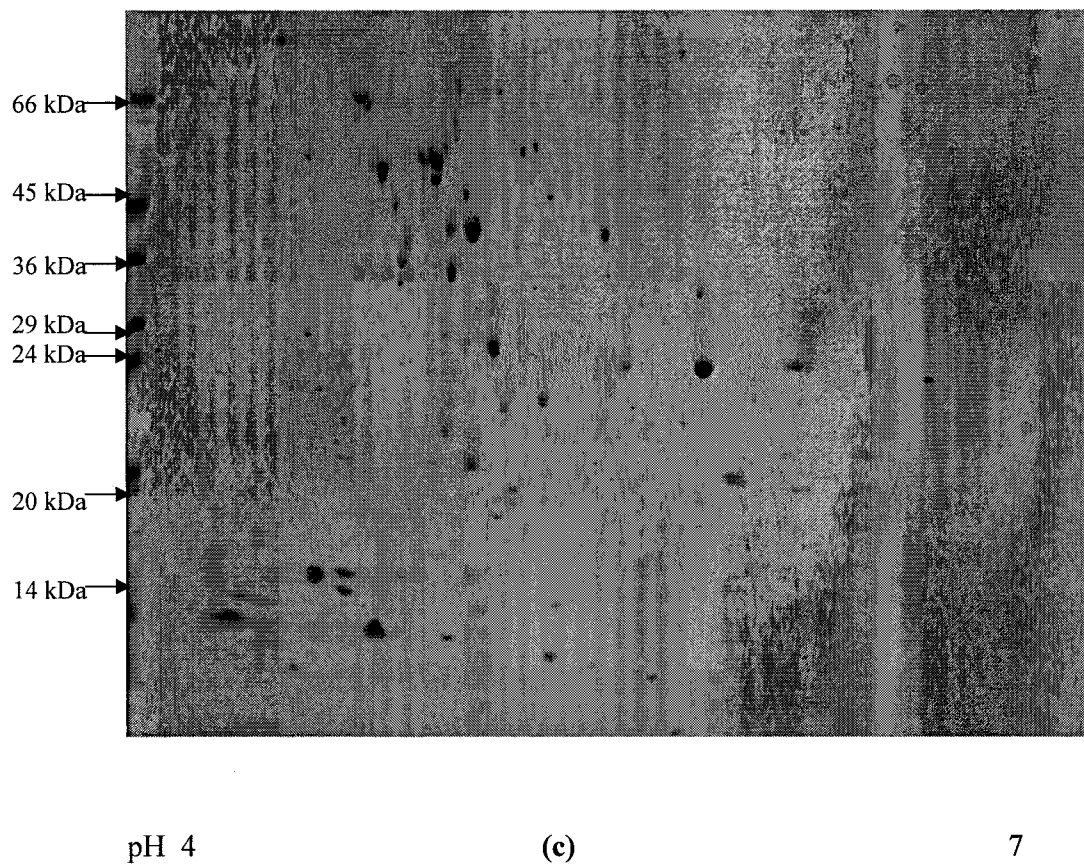
(a)

7

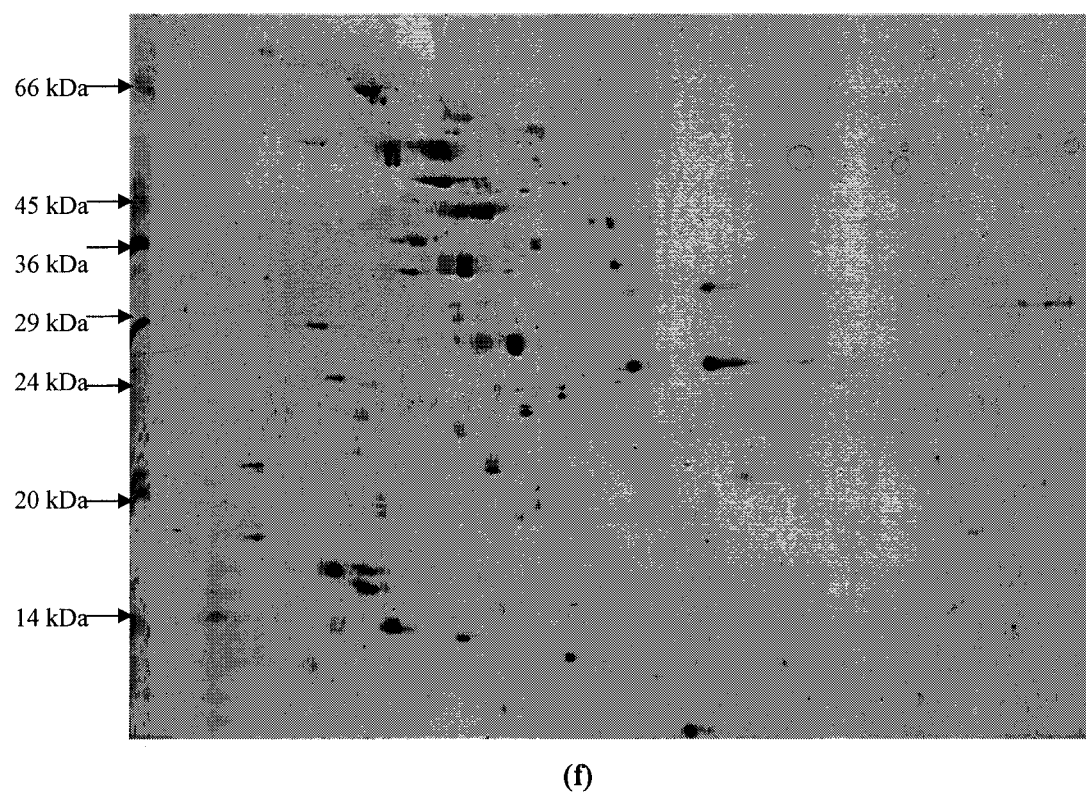
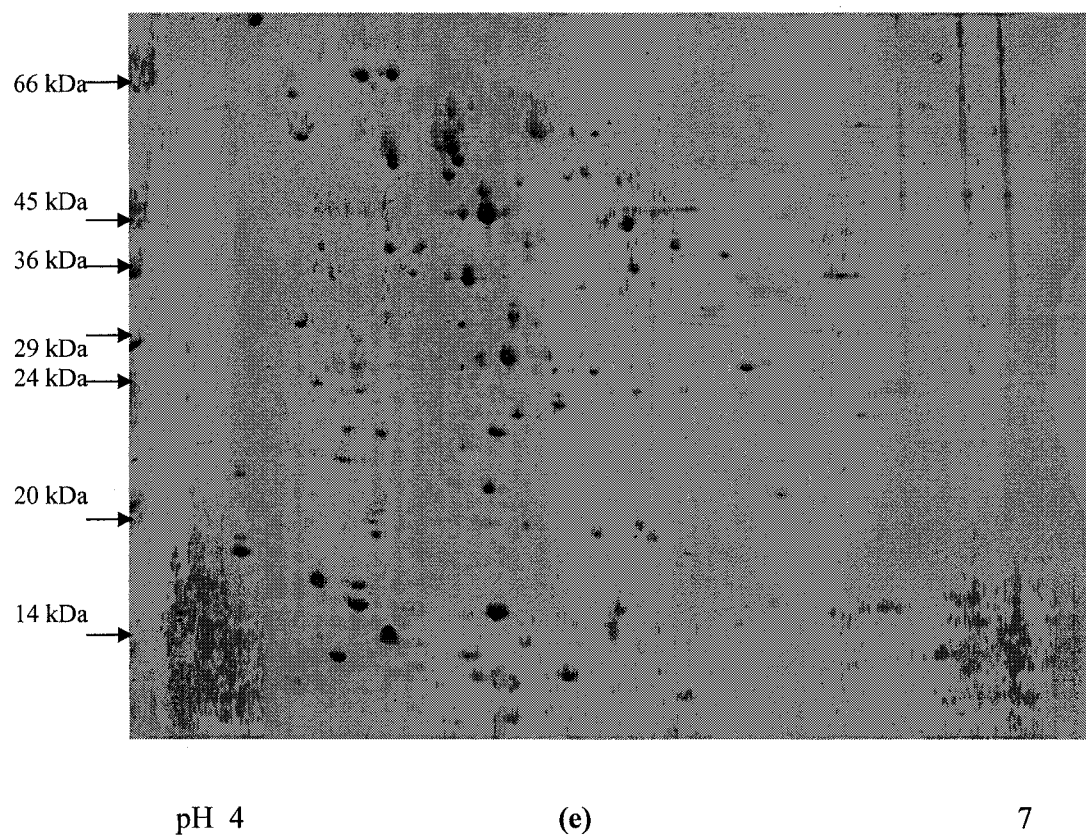


(b)

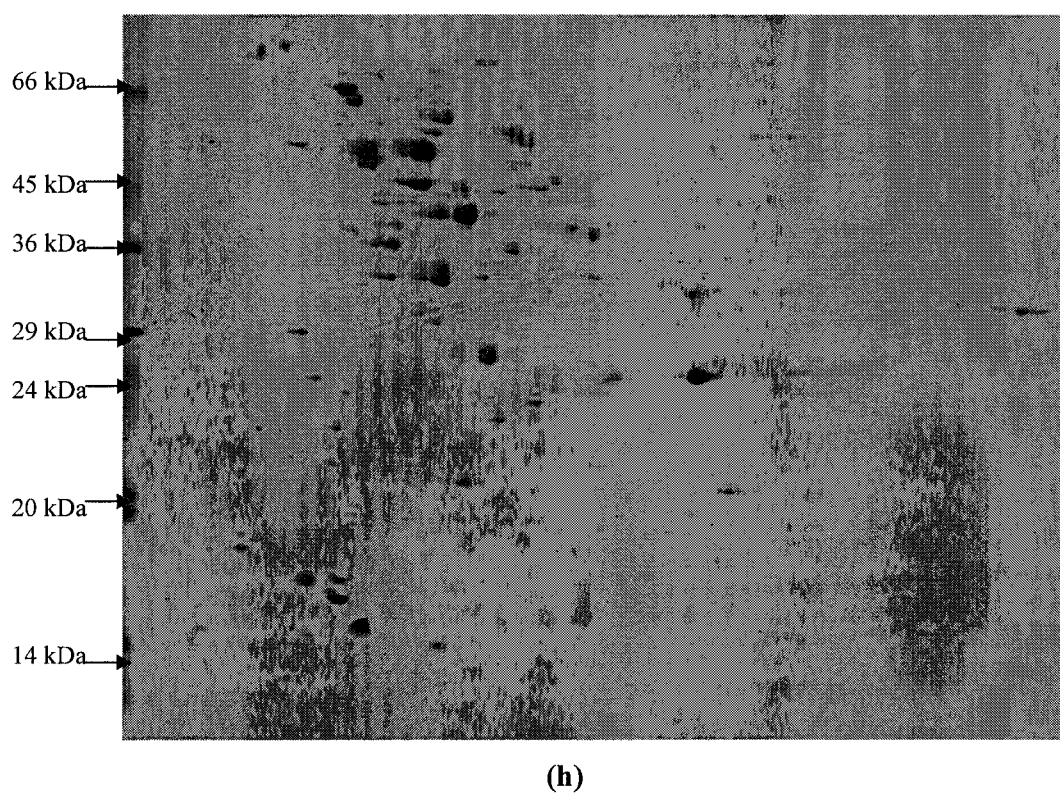
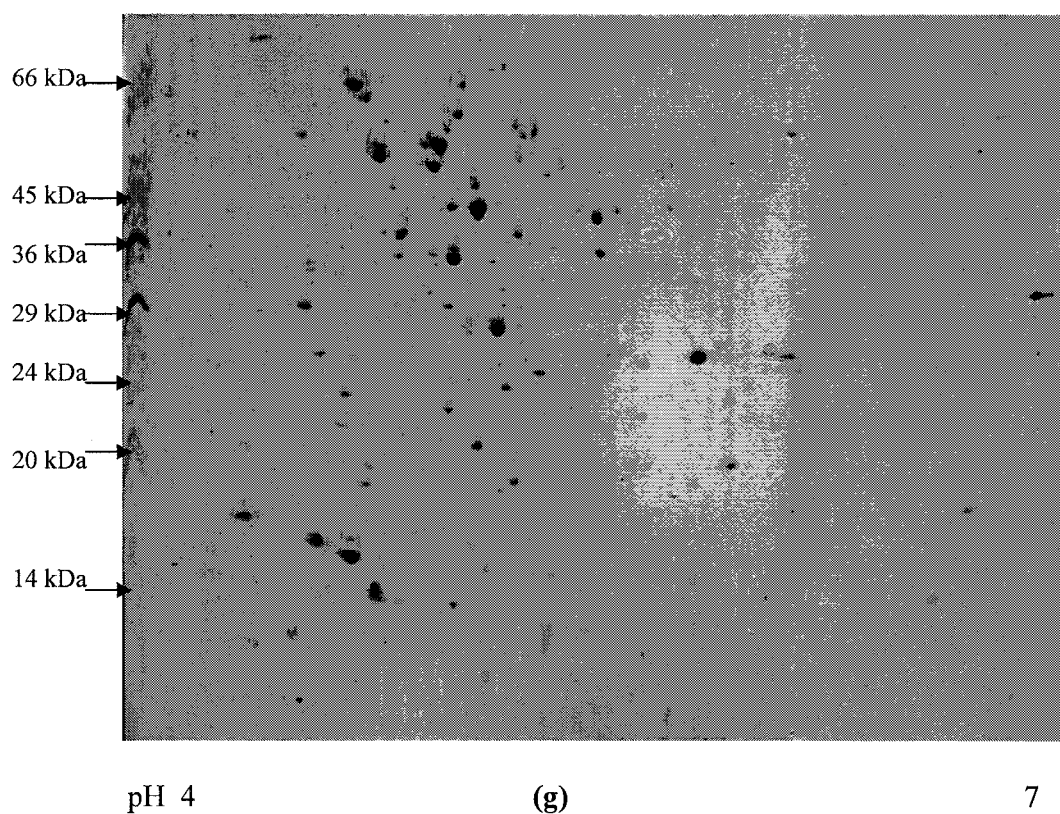
(Figure continues on next page)



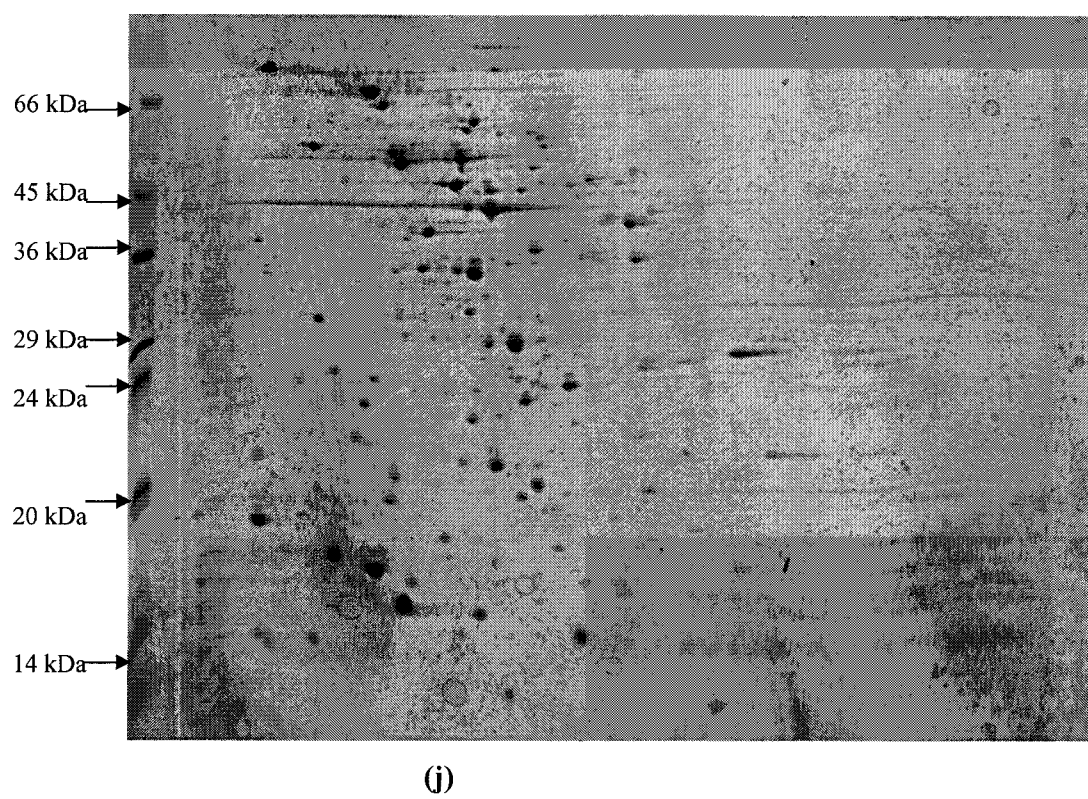
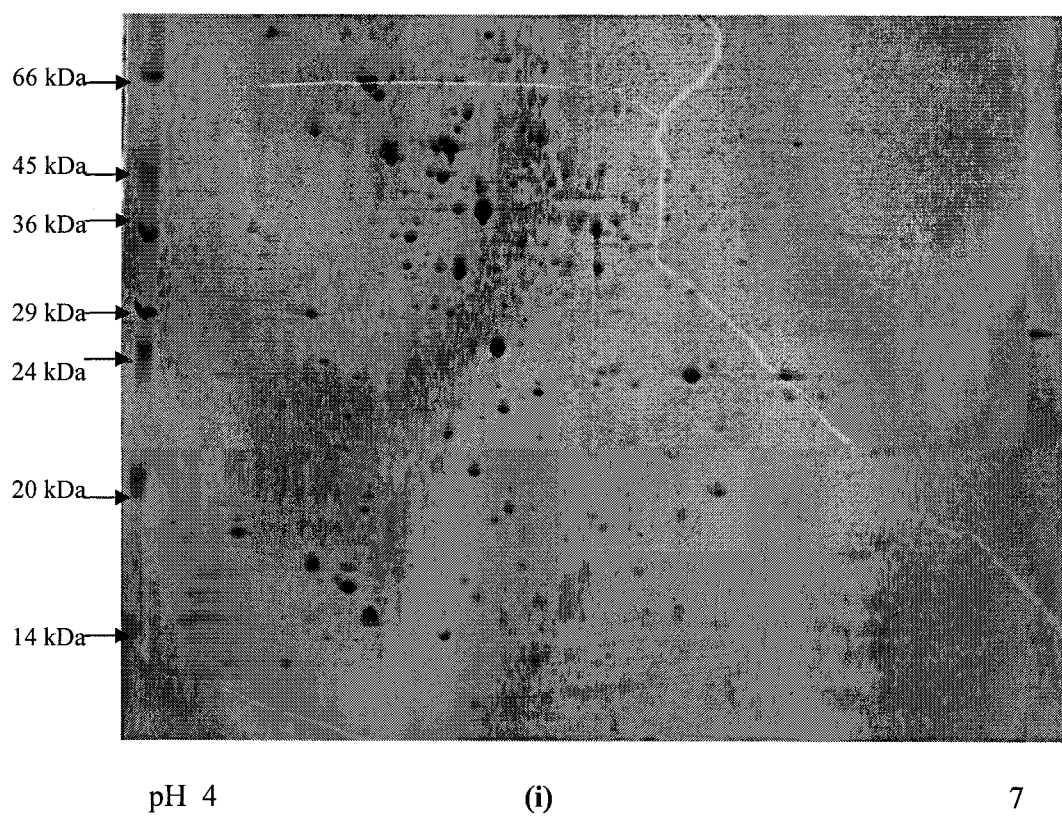
(Figure continues on next page)



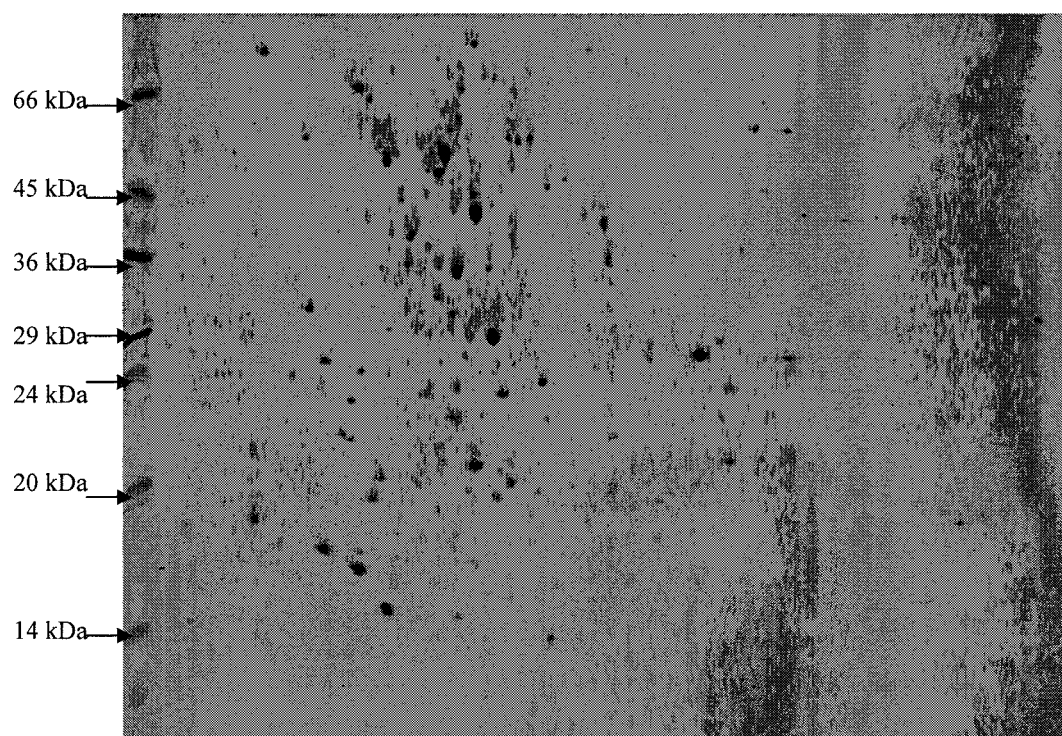
(Figure continues on next page)



(Figure continues on next page)



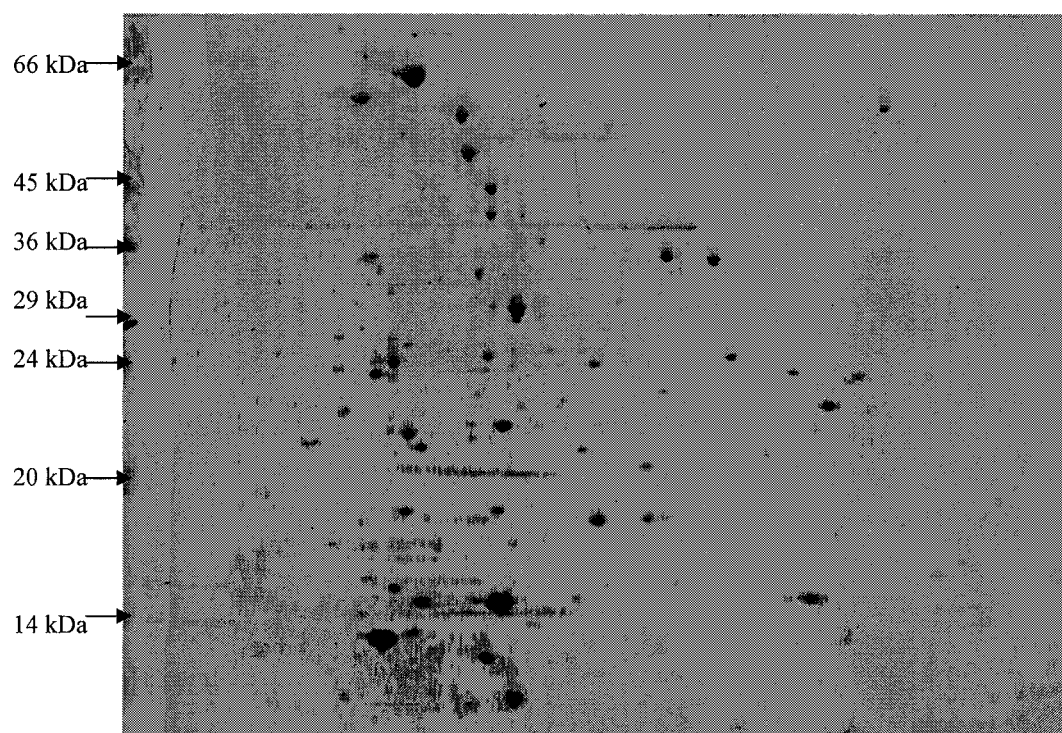
(Figure continues on next page)



pH 4

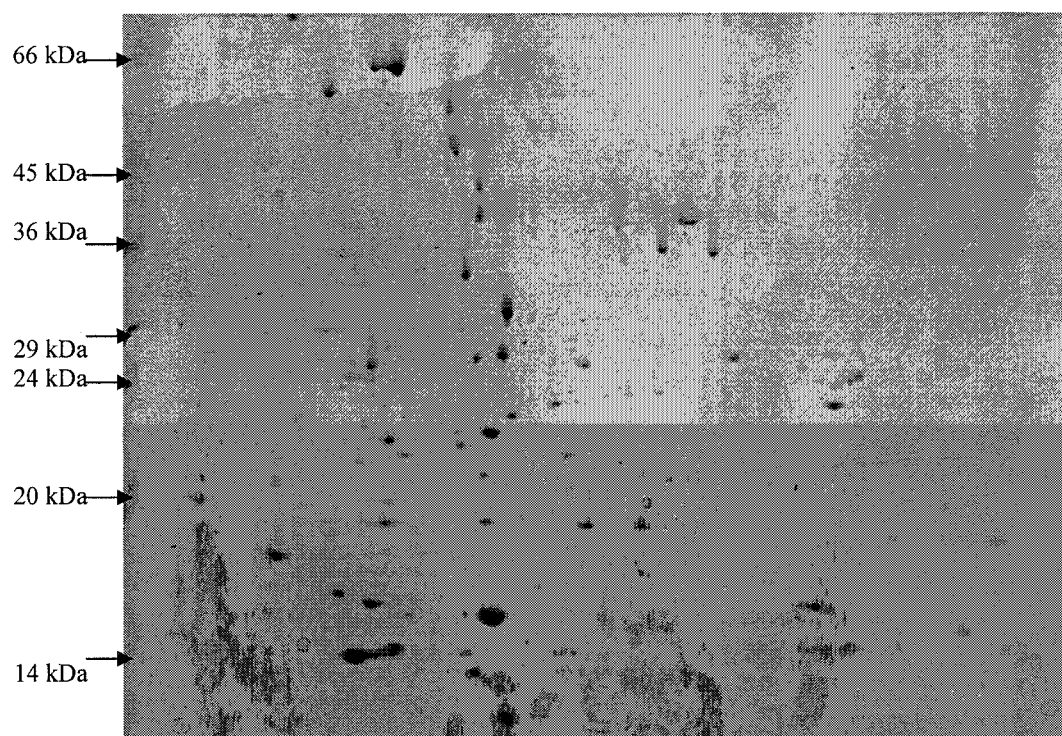
(k)

7



(l)

(Figure continues on next page)



pH 4

(m)

7



(n)

(Figure continues on next page)

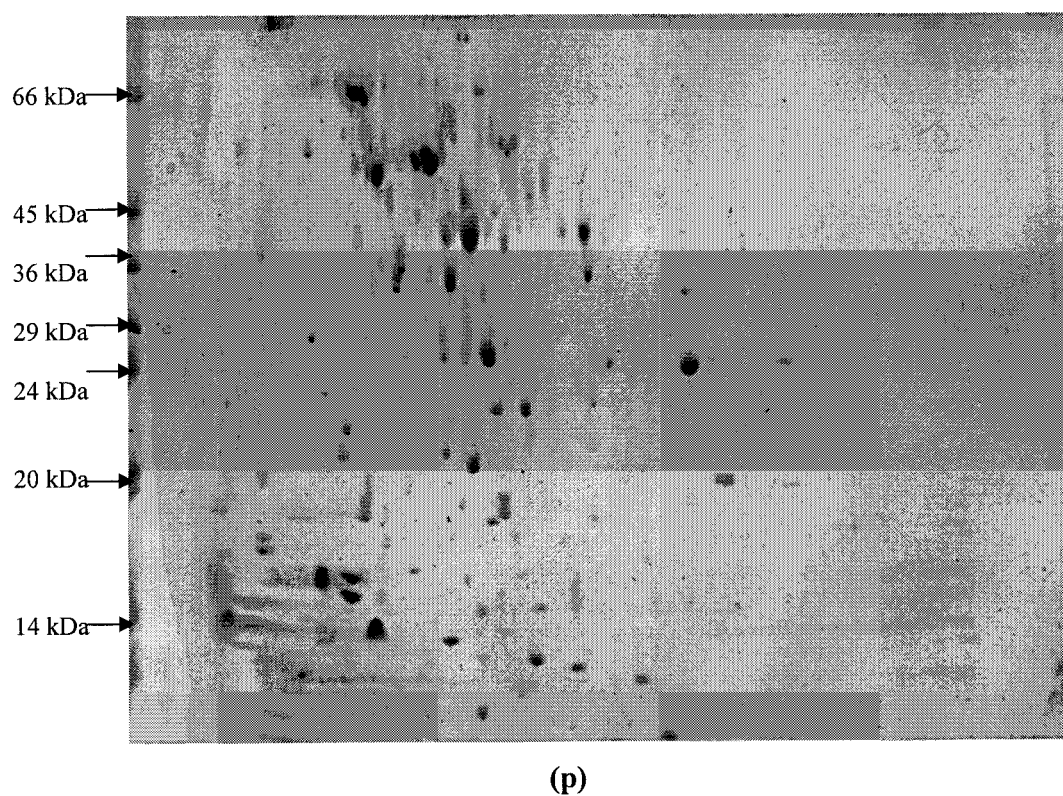
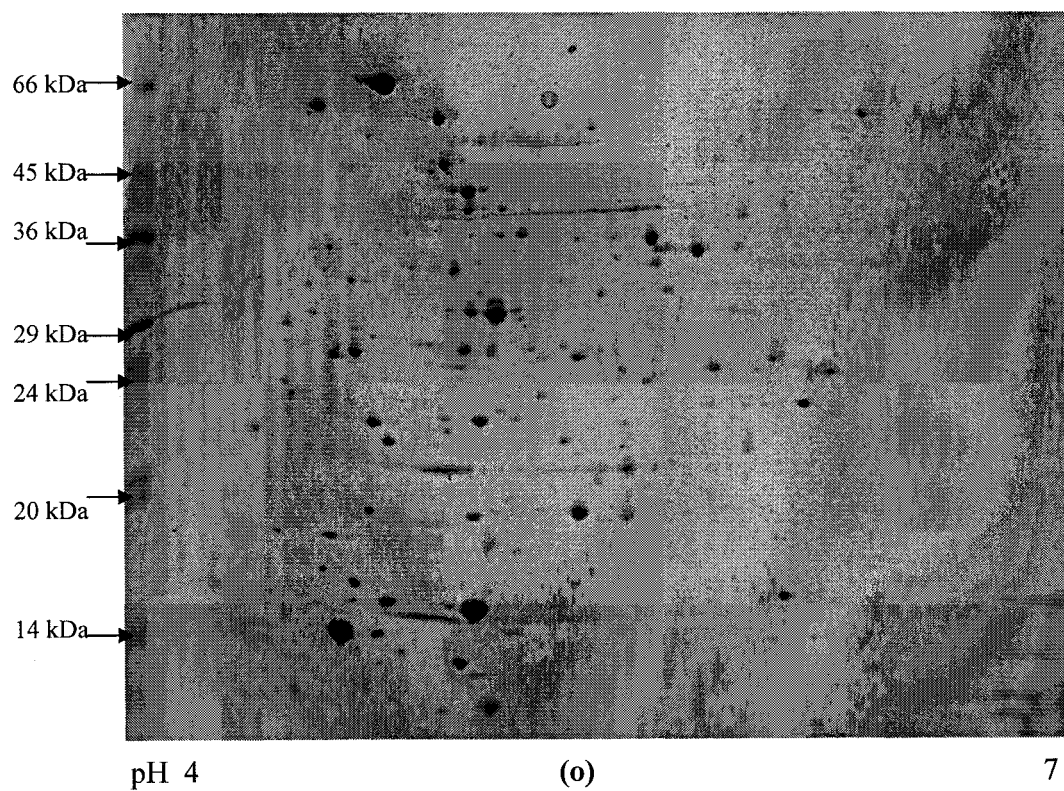


Figure 4.3 Comparison of cell-associated proteins from *S. pyogenes* SF370.

Proteome maps for *S. pyogenes* SF370 growing in HA-rich media (a, c, e, g, i, k, m, o) and control media (b, d, f, h, j, l, n, p).

4.2.7 Identification of GAS proteins from 2DE gels using static nanospray LCQ Advantage MS with BioWorks 3.2 software

The protein spots corresponding to the 18 differentially expressed proteins (Figure 4.4) were excised from gels and subjected to in-gel trypsin digestion. The resulting peptides were analysed by static nanospray LCQ Advantage MS. The spectra of 18 protein digests were opened in BioWorks 3.2 software. The MS/MS data was searched against the streptococci database using TurboSEQUEST search function. Of the 18 proteins analysed spots 1-11 were identified according to the filter parameters in section 2.10.1.3. (Figure 4.5 - Figure 4.15). The MS/MS profile for each identified peptide is shown in Appendix E

The remaining seven proteins could not be added to the list of identified proteins as they did not satisfy the filter parameters defined in the section 2.10.1.3. Among these seven spots, spots 14, 16 and 17 gave poor quality MS data and therefore could not be analysed further. However, spots 12 (Figure 4.16), 13 (Figure 4.17), 15 (Figure 4.18) and 18 (Figure 4.19) gave good quality of MS data.

4.2.8 Protein identification and characterisation using TSSView™ v 1.0

4.2.8.1 Protein identification via highly-conserved sequence identification

Analysis of MS data of protein spots 12, 13, 15 and 18 provided high quality peptides (Figure 4.16, 4.17, 4.18 and 4.19), however they were not added to the list of identified proteins because either the peptides are not from SF370 strain, or the peptides are from SF370 strain but only one peptide was identified. For further identification, these peptides were analysed using detailed bioinformatics screening using TSSView™ v 1.0.

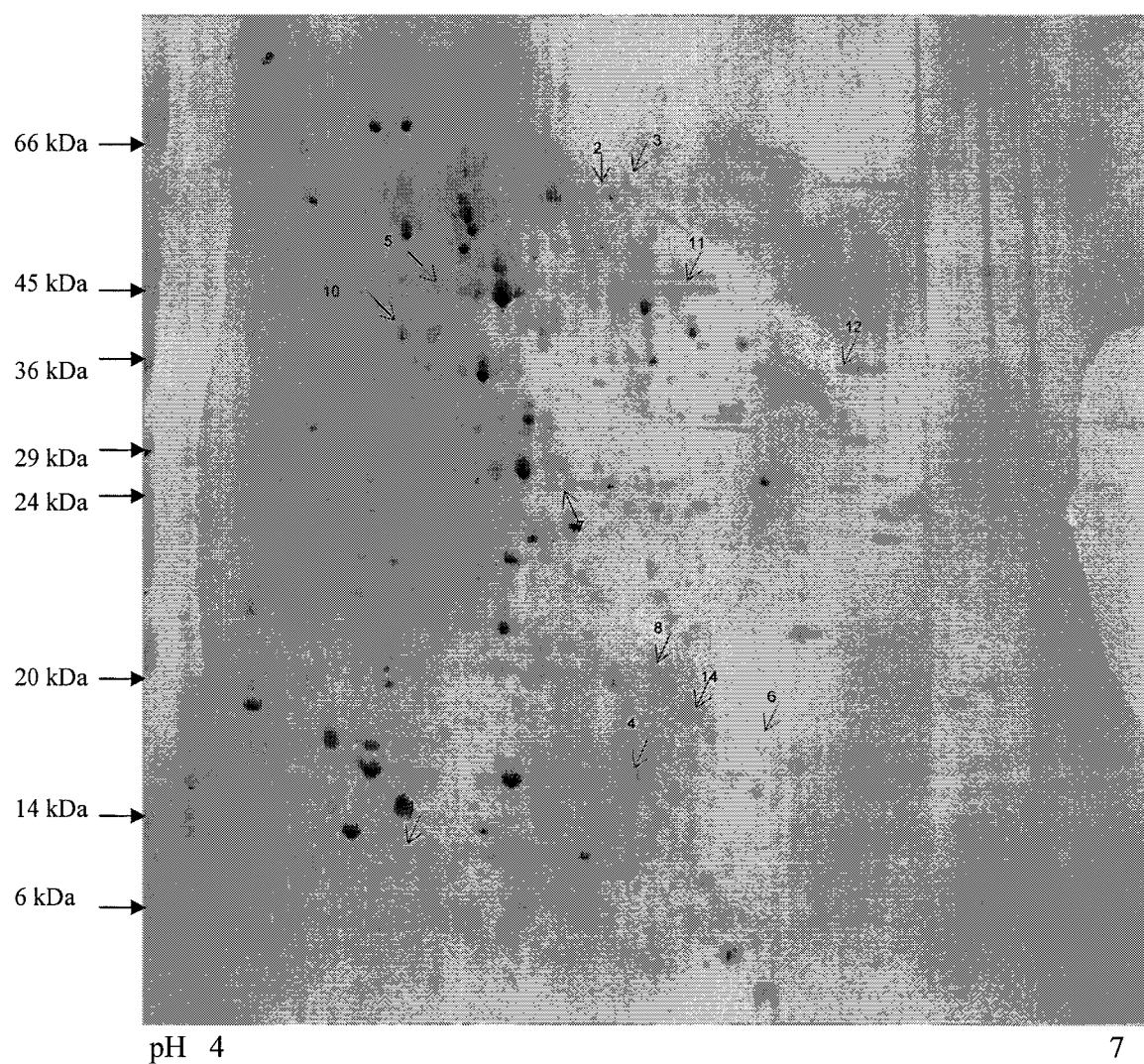


Figure 4.4 (a) 2DE gel analysis of cell-associated proteins from GAS.

Cells cultured in HA-rich media; the numbers indicate those proteins that were up-regulated in HA-rich media.

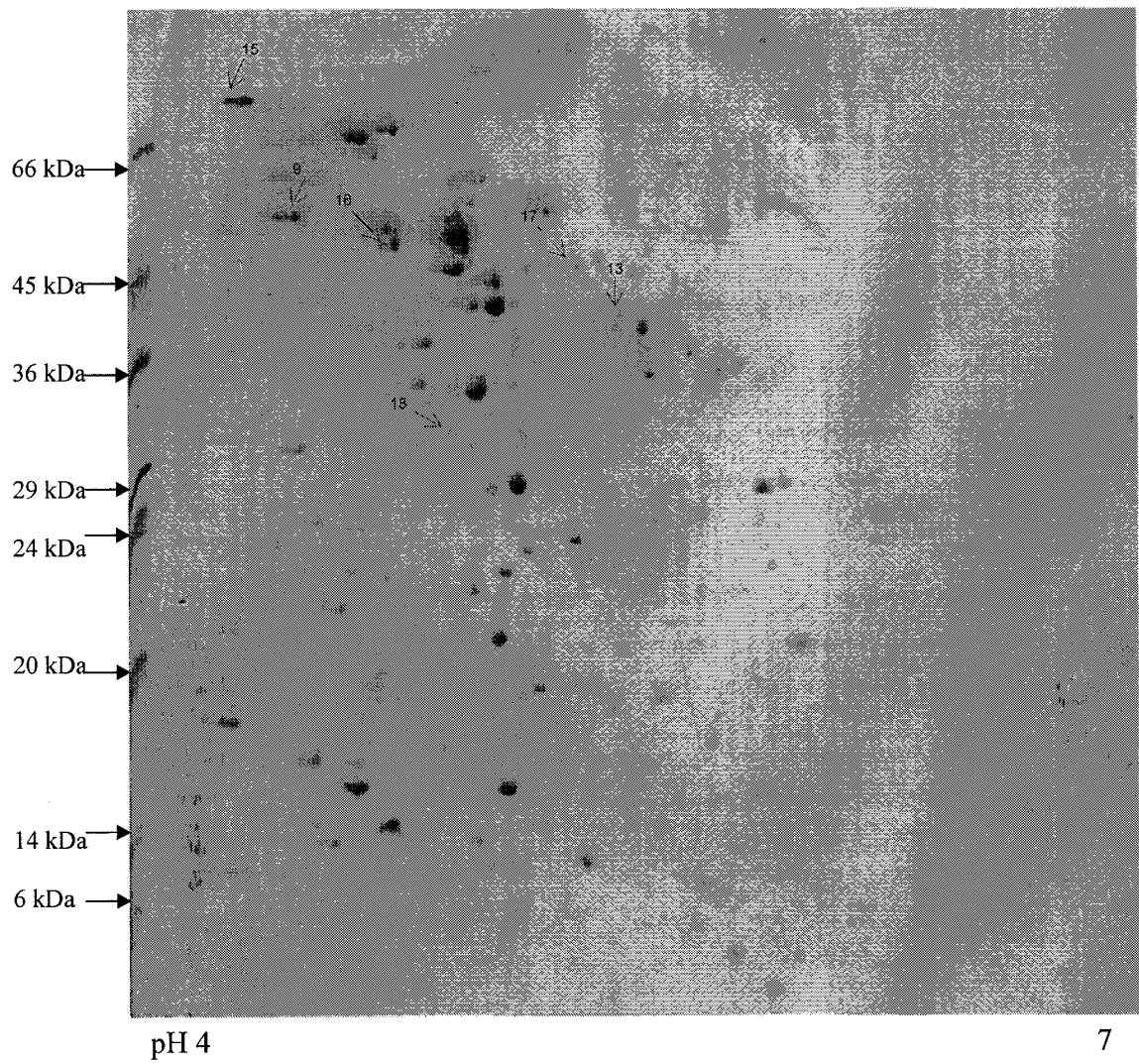
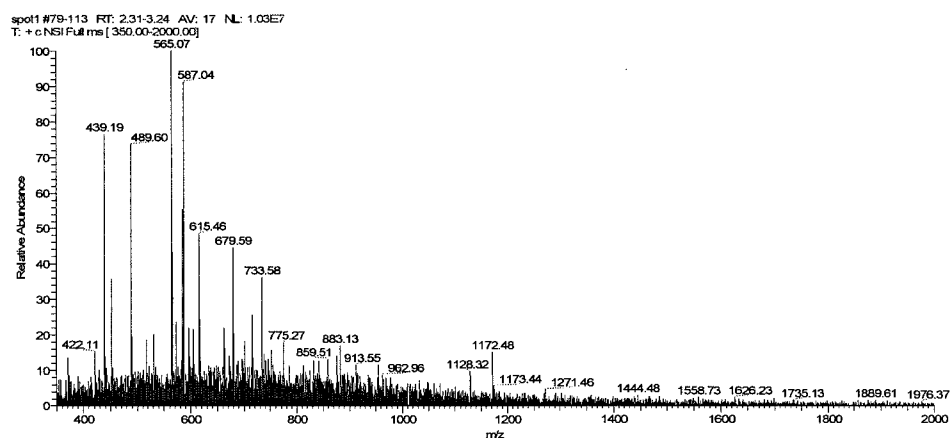


Figure 4.4 (b) 2DE gel analysis of cell-associated proteins from GAS.

Cells cultured control media; the numbers indicate those proteins that were down-regulated in HA-rich media.



Mass spectrum of protein digestion

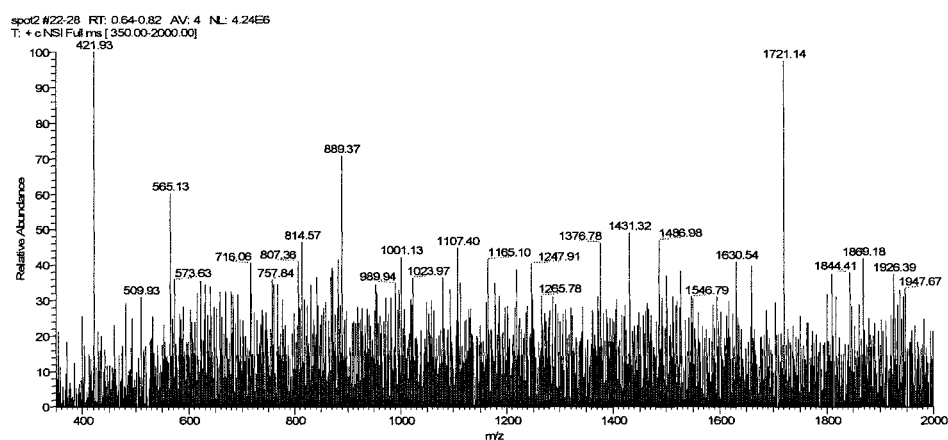
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Putative phosphotransferase system phosphohistidine- containing protein	K.FASDITLDYK.G	2	3.198	0.694
	R.PATLLVQTASK.F	2	3.814	0.602
	K.FASDITLDYKKGK.A	3	2.768	0.549
	K.FASDITLDYKKGK.A	2	3.714	0.707
	R.PATLLVQTASK.F	1	2.189	0.541

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
PATLLVQTASK	1129.33	12.64	18 - 28	12.64
FASDITLDYK	1173.30	13.13	29 - 38	11.49
FASDITLDYKKGK	1358.52	15.20	29 - 40	13.79
Total		27.61		26.44

Protein coverage by matched peptides

Figure 4.5 Identification of protein spot No.1 as a putative phosphotransferase system phosphohistidine-containing protein (HPr).



Mass spectrum of protein digestion

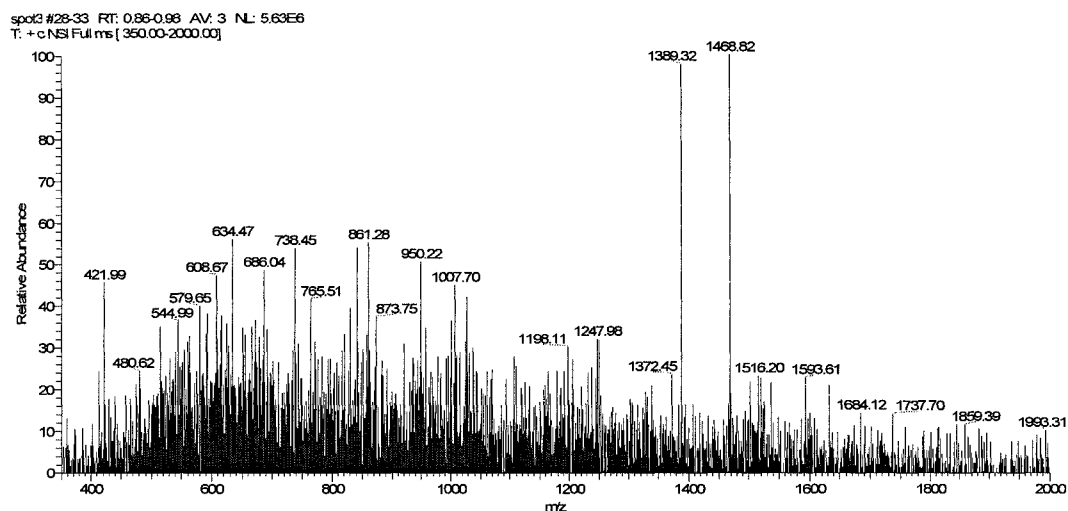
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	R.DAEKIGAILSK.E	2	0.856	0.120
	K.NIVAGAFGYSGQR.C	2	2.847	0.685
	K.AFGIAEQLEVGTVHLNNK.T	2	2.043	0.348

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
DAEKIGAILSK	1145.33	2.27	78 - 88	2.32
NIVAGAFGYSGQR	1340.47	2.66	271 - 283	2.74
AFGIAEQLEVGTVHLNNK	1941.18	3.85	417 - 434	3.79
Total		8.71		8.84

Protein coverage by matched peptides

Figure 4.6 Identification of protein spot No.2 as a putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase



Mass spectrum of protein digestion

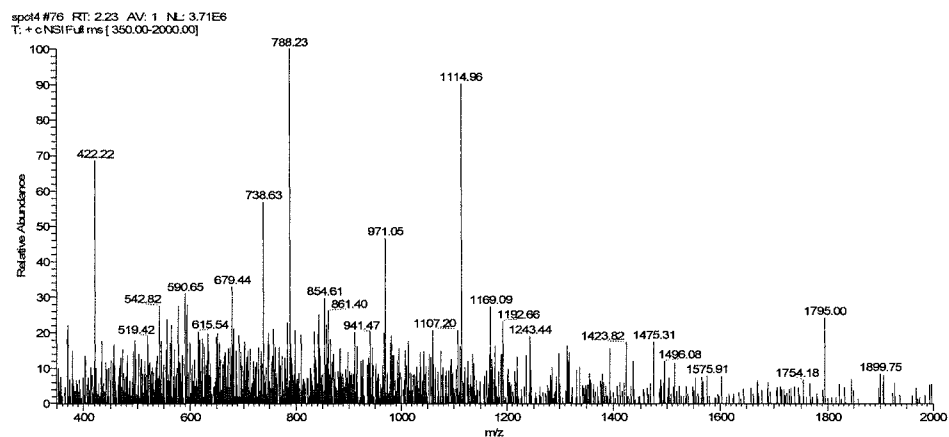
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
M1 protein precursor	K.ALELAIDQASQDYNR.A	2	1.245	0.338
	K.ARLENAMVAGRDFK.R	2	1.118	0.366

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
ALELAIDQASQDYNR	1707.82	6.57	75 - 89	6.61
ARLENAMVAGRDFK	1707.94	6.57	171 - 185	6.61
Total		13.07		13.22

Protein coverage by matched peptides

Figure 4.7 Identification of protein spot No.3 as a M1 protein precursor



Mass spectrum of protein digestion

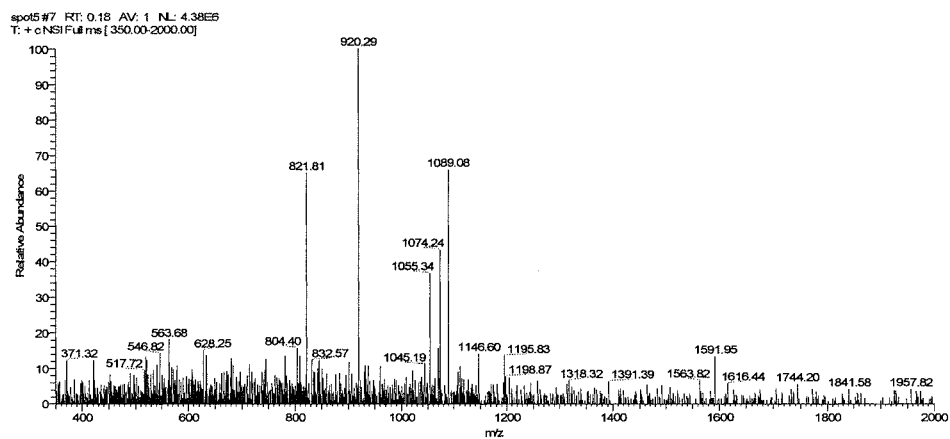
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Hypothetical protein SPy1143	R.AVILRQGAIK.E	2	1.535	0.255
	K.TGPLIGPSANLSGKASGR.V	2	1.603	0.428

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
TGPLIGPSANLSGKASGR	1683.89	7.79	120 - 137	9.18
AVILRQGAIK	1170.43	5.41	174 - 184	5.61
Total		13.11		14.80

Protein coverage by matched peptides

Figure 4.8 Identification of protein spot No.4 as a hypothetical protein SPy1143



Mass spectrum of protein digestion

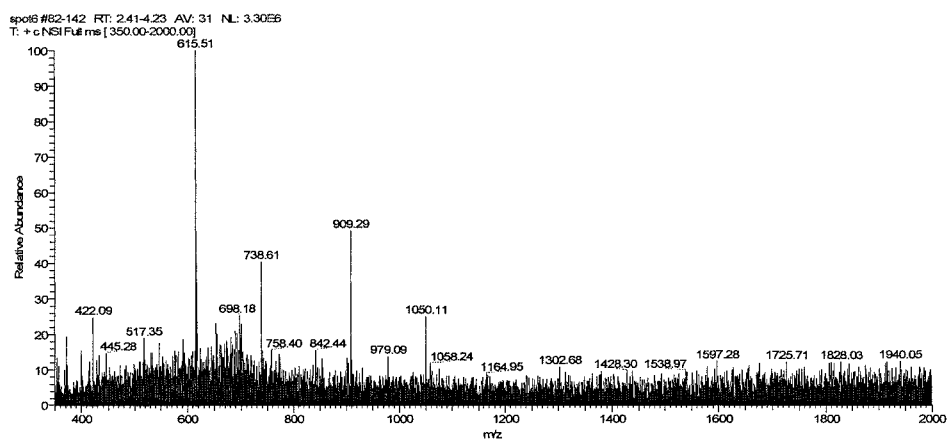
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Histidine protein kinase	K.DLVFLLRHFEK.A	2	1.000	0.581
	R.HFEKAVRNPLAHLIK.P	2	1.000	0.664
	R.REPFYFDDMNAIK.K	2	0.527	0.357

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
DLVFLLRHFEK	1417.68	4.60	182 - 192	4.20
HFEKAVRNPLAHLIK	1774.11	5.76	189 - 203	5.73
REPFYFDDMNAIK	1760.01	5.71	238 - 251	5.34
Total		14.18		13.74

Protein coverage by matched peptides

Figure 4.9 Identification of protein spot No.5 as a histidine protein kinase



Mass spectrum of protein digestion

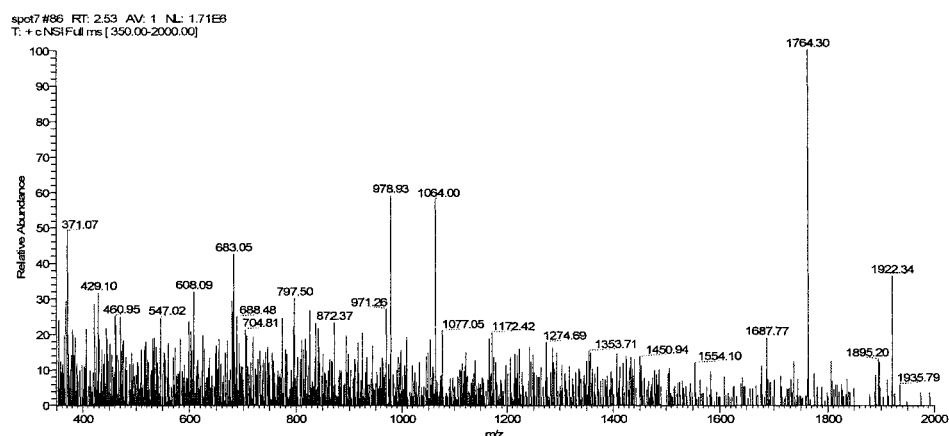
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
tRNA modification GTPase	R.VDLTQAEAVMDIIR.A	2	1.945	0.250
	K.TGTIIDEVMVSVMLAPK.T	3	2.047	0.248

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
TGTIIDEVMVSVMLAPK	1805.19	3.57	63 - 79	3.71
VDLTQAEAVMDIIR	1574.82	3.12	128 - 141	3.06
Total		6.65		6.77

Protein coverage by matched peptides

Figure 4.10 Identification of protein spot No.6 as a tRNA modification GTPase



Mass spectrum of protein digestion

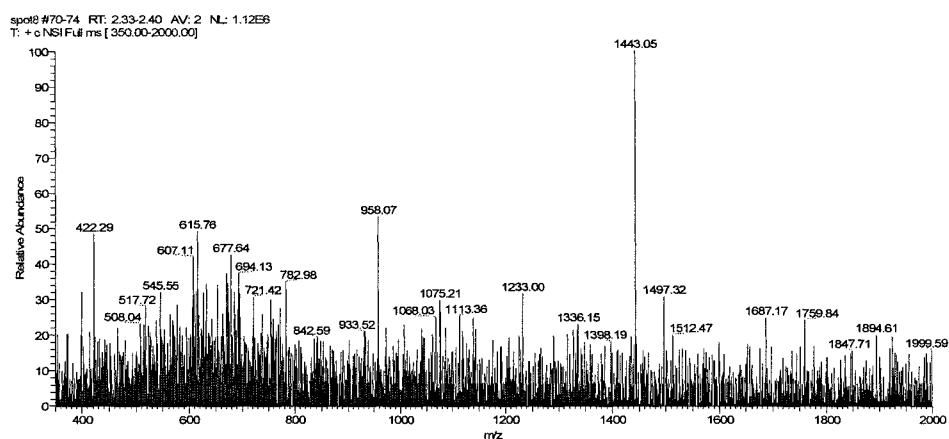
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Hypothetical protein SPy1262	R.MTDLDVIEVNVK.V	2	4.304	0.600
	K.LVNTESVR.D	2	1.882	0.480
	R.DGVNVEVGKK.Q	2	1.199	0.113
	R.GQLTYDDK.V	2	2.516	0.496
	R.GQLTYDDKVIEK.I	3	1.538	0.204
	K.SIVEEEVKR.M	2	1.292	0.433
	R.STSEFTSHQVENVK.A	2	3.243	0.557
	K.ASVDNGVEK.L	2	1.207	0.187
	K.DL TSAIR.G	2	0.648	0.315

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
DL TSAIR	775.87	3.89	12 - 18	3.91
GQLTYDDK	939.99	4.71	19 - 26	4.47
GQLTYDDKVIEK	1409.57	7.07	19 - 30	6.70
LVNTESVR	918.03	4.60	57 - 64	4.47
DGVNVEVGKK	1045.17	5.24	65 - 74	5.59
SIVEEEVKR	1089.22	5.46	99 - 107	5.03
MTDLDVIEVNVK	1376.60	6.90	108 - 119	6.70
STSEFTSHQVENVK	1593.68	7.99	146 - 159	7.82
ASVDNGVEK	918.97	4.61	160 - 168	5.03
Total		45.07		45.25

Protein coverage by matched peptides

Figure 4.11 Identification of protein spot No.7 as a hypothetical protein SPy1262



Mass spectrum of protein digestion

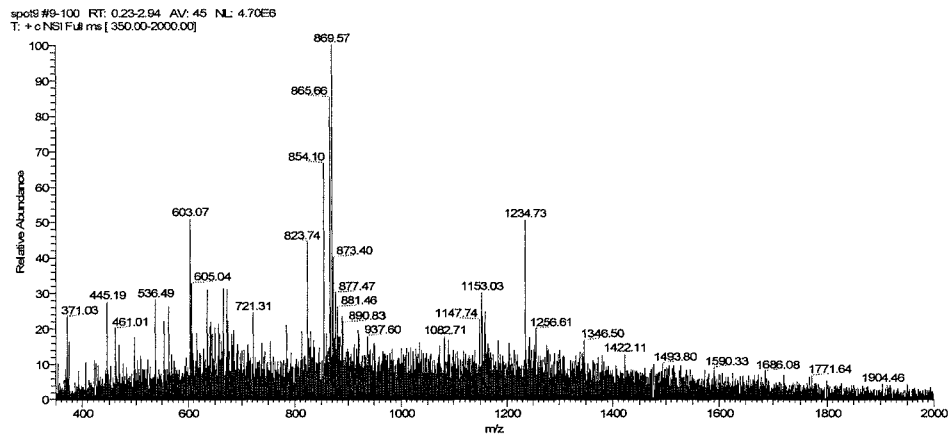
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
50S ribosomal protein L10	K.KAEQVELIAEK.M	2	4.582	0.565
	R.GLTVDQDTVLR.R	2	2.122	0.422
	K.AAASIVIVDSR.G	2	3.835	0.693
	R.SLRESGVEFK.V	2	2.052	0.417
	R.GLTVDQDTVLR.R	2	0.523	0.102
	K.VINDFTK.T	1	1.181	0.455

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
KAEQVELIAEK	1258.45	7.16	9 - 19	6.63
AAASIVIVDSR	1102.27	6.27	22 - 32	6.63
GLTVDQDTVLR	1217.35	6.93	33 - 43	6.63
GLTVDQDTVLR.R	1373.54	7.83	33 - 44	7.23
SLRESGVEFK	1152.28	6.56	45 - 54	6.02
VINDFTK	836.96	4.76	96 - 102	4.22
Total		27.61		30.72

Protein coverage by matched peptides

Figure 4.12 Identification of protein spot No.8 as a 50S ribosomal protein L10



Mass spectrum of protein digestion

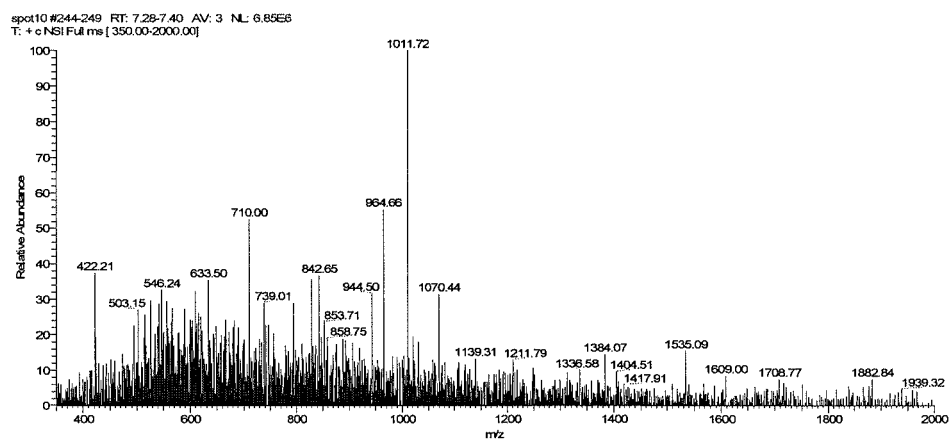
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
RopA	K.AVEVITSTASVK	2	2.313	0.645
	R.VKTNLVIEAIAK.A	2	2.409	0.574
	K.GKEWTLSAEVVTKPEVK.L	3	4.790	0.570
	R.GVITFTISQDK.I	2	1.750	0.523
	K.TNLVIEAIAK.A	2	2.674	0.361

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
GVITFTISQDK	1209.37	2.57	13 - 23	2.58
GKEWTLSAEVVTKPEVK	1902.18	4.04	101 - 117	3.98
VKTNLVIEAIAK	1299.58	2.76	358 - 369	2.81
TNLVIEAIAK	1072.28	2.28	360 - 369	2.34
AVEVITSTASVK	1205.38	2.56	416 - 427	2.81
Total		11.80		12.18

Protein coverage by matched peptides

Figure 4.13 Identification of protein spot No.9 as a RopA



Mass spectrum of protein digestion

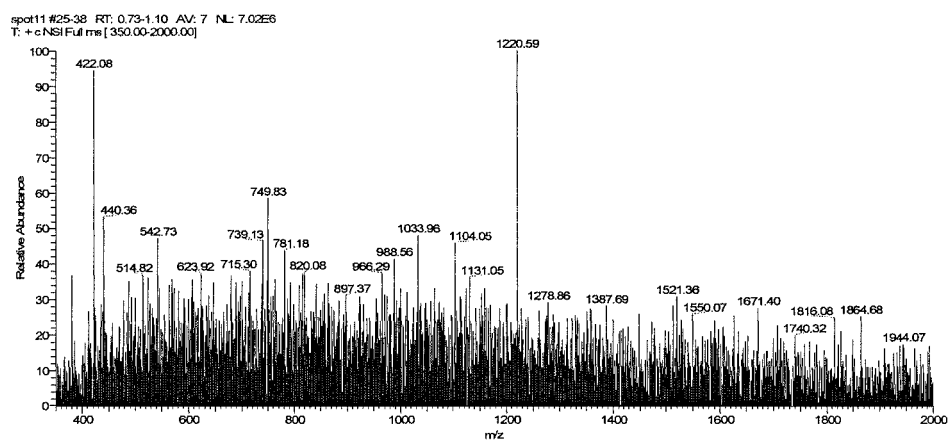
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Collagen-like surface protein	R.GETFAQGPVGPQGEK.G	2	1.490	0.222
	R.GEQGIQ GKAGEKGER.G	2	1.604	0.229

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
GEQGIQ GKAGEKGER	1544.65	4.26	142 - 159	4.31
GETFAQGPVGPQGEK	1412.49	3.90	205 - 219	4.31
Total		8.10		8.62

Protein coverage by matched peptides

Figure 4.14 Identification of protein spot No.10 as a collagen-like surface protein.



Mass spectrum of protein digestion

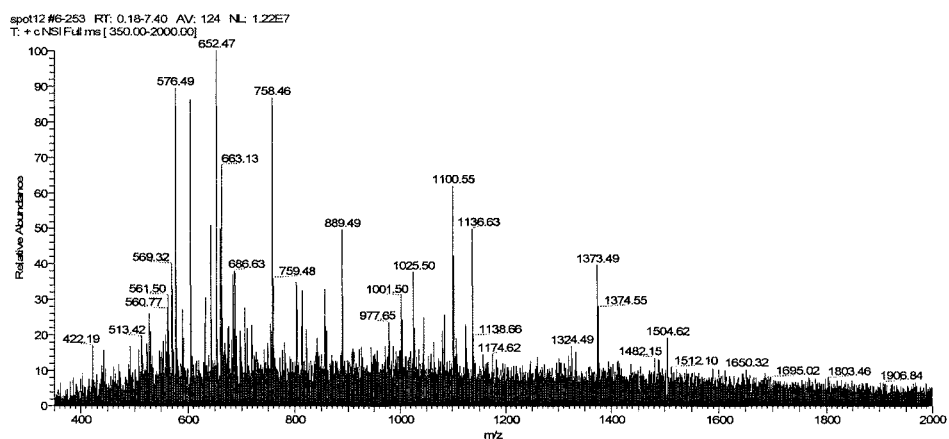
Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Putative carbamoyl-phosphates synthase small subunit	K.GIPGISGIDTRALTK.I	2	1.439	0.252
	K.MTFGHRGFNHAVR.E	1	1.330	0.275

Information of matched peptides

Sequence	MH+	% by Mass	Position	% by AA's
GIPGISGIDTRALTK	1499.74	3.77	105 - 119	4.17
MTFGHRGFNHAVR	1530.74	3.85	264 - 276	3.61
Total		7.57		7.78

Protein coverage by matched peptides

Figure 4.15 Identification of protein spot No.11 as a putative carbamoyl-phosphates synthase small subunit

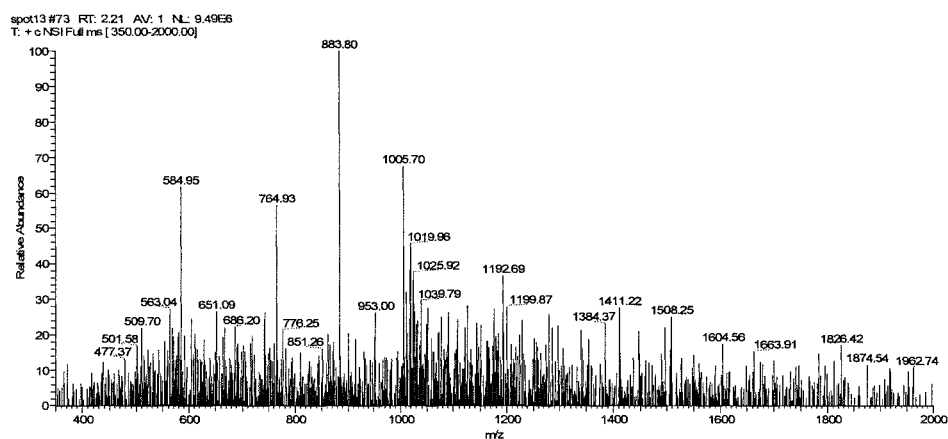


Mass spectrum of protein digestion

Protein	Peptides sequence	Charge	Xcorr	ΔC_n
transcriptional repressor CodY	K.DTYPGGLTTIPIYGGGMRL	2	1.826	0.361

Information of matched peptide

Figure 4.16 Identification of protein spot No.12 as a transcriptional repressor CodY from *S. pyogenes* SF370

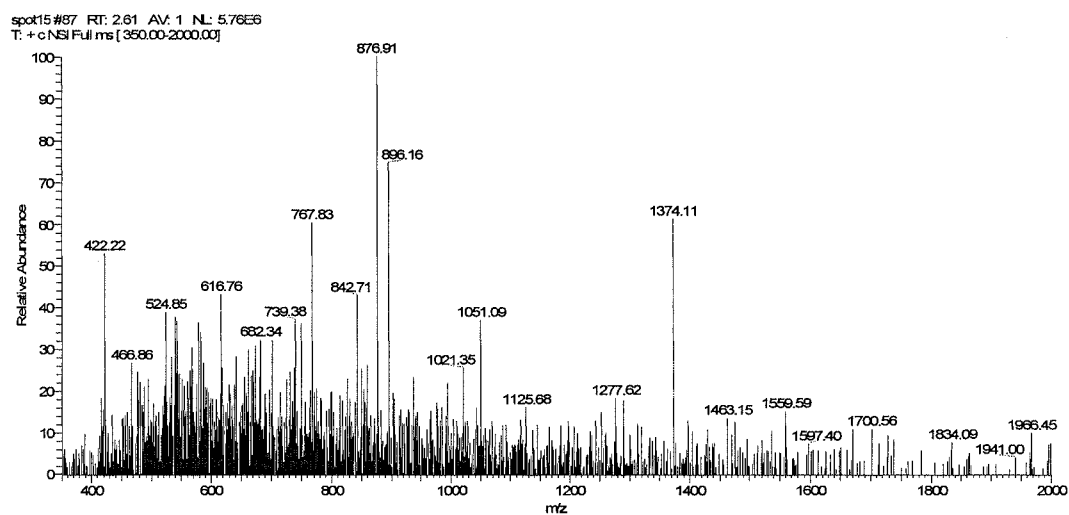


Mass spectrum of protein digestion

Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Putative formate-tetrahydrofolate ligase	K.LVLVTAINPTPAGEGK.S	3	1.685	0.541
	K.STITIGLADALNKIGK.K	2	1.484	0.392

Information of matched peptides

Figure 4.17 Identification of protein spot No.13 as a formate-tetrahydrofolate ligase from *Streptococcus mutans* UA159

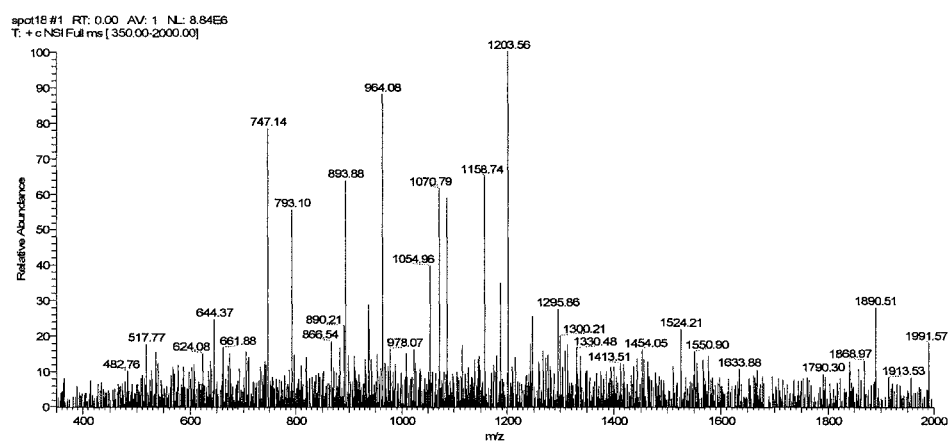


Mass spectrum of protein digestion

Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Alcohol dehydrogenase, zinc-containing	K.LSEIEKAYEVFSK.A	2	1.935	0.472

Information of matched peptide

Figure 4.18 Identification of protein spot No.15 as a zinc-containing alcohol dedehydrogenase from *S. pneumoniae* TIGR4



Mass spectrum of protein digestion

Protein	Peptides sequence	Charge	Xcorr	ΔC_n
Unknown protein	K.SSDSSVGEETLPSPSLK.S	2	1.555	0.307
	R. TKRAVPGEPATPDK.K	2	1.606	0.419

Information of matched peptides

Figure 4.19 Identification of protein spot No.18 as an unknown protein from *S. pneumoniae*

Protein spot 12, that did not satisfy the filter parameters as only one peptide was successfully matched, was identified as transcription repressor CodY, as the successfully matched peptide was shown using TSSview™ to come from a highly-conserved region of the transcription repressor CodY (O/E ratio: ∞). This single peptide is therefore highly significant for assigning function and therefore facilitates protein identification. Protein spot 13, that did not satisfy the filter parameters as it is from *Streptococcus mutans* and not *Streptococcus pyogenes*, was identified as a putative formate-tetrahydrofolate ligase as the two successfully matched peptides were shown using TSSview™ to come from highly-conserved regions of a putative formatetetrahydrofolate ligase (O/E ratio: ∞). A BlastPAM30 search of these peptides confirmed that highly-identical peptides are present in the same protein from the SF370 strain of the M1 GAS serotype (Table 4.3).

The remaining two proteins (spots 15 and 18) could not be identified using TSSview™ and BlastPAM30 because the peptides from these proteins are located in conserved regions of proteins that are not present in any homologues of the SF370 strain of the M1 serotype of GAS.

4.2.8.2 Functional analysis of hypothetical proteins

From 18 protein spots, 13 proteins were identified using BioWorks 3.2 software and/or TSSView™ v1.0 software, including the two hypothetical proteins, SPy1143 and SPy1262, the functions of which have not been annotated in the genome database. The potential functions of these two proteins were analysed using TSSView™ v1.0 software by searching the full length sequences of the two proteins. Several conserved regions for SUA5 functionality were found in hypothetical protein SPy1143 (Figure 4.20) and conserved regions for gls24 (Figure 4.21), a general stress protein, were found in hypothetical protein, SPy1262.

Peptide sequence	BlastPAM30 alignment	E value
LVLVTAINPTPAGEGK	Query: 1 LVLVTAINPTPAGEGK 16 L+LVTAINPTPAGEGK Sbjct: 56 LILVTAINPTPAGEGK 71	1.0×10^{-5}
STITIGLADALNKIGKK	Query: 1 STITIGLADALNKIGKK 17 ST IGLADALN GKK Sbjct: 72 STMSIGLADALNQMGKK 88	1.4×10^{-2}

Table 4.3 BlastPAM30 alignments for two peptides of the putative formatetetrahydrofolate ligase.

Identified by analysing MS data of protein spot 13.

4.2.8.3 Functional characterisation of identified proteins

All the other identified proteins were also analysed using TSSView™ v1.0 for further functional characterisation. In addition to their genome-sequence annotated function, the regions of unusually high conservation in the sequences were also bioinformatically characterized. The M protein sequence contains the conserved region for plasminogen-binding activity, conserved region for fibrinogen binding activity, and conserved region for Fc- γ - receptor (Figure 4.22). These activities of highly-conserved regions were therefore assigned as additional potential function of M protein. In the sequence of tRNA modification GTPase, the conserved regions for thiophene and furan oxidation protein activity, ThdF protein-activity, GTP-binding protein engA activity, tRNA synthase activity and cytidylate kinase activity were identified (Figure 4.23).

4.2.9 Comparison of differentially expressed secreted proteins of *S. pyogenes* SF370 growing in HA-rich media and control media using 2DE

The secreted proteins of *S. pyogenes* SF370 from supernatant in control media and HA-rich media were isolated and separated using the techniques optimized in section 3.2.2.1.3. Two pairs of 2DE gels were constructed; one in pH range 3 – 10 (Figure 4.24) and another in pH range 6 – 11 (Figure 4.25). Proteins ($100 \pm 10 \mu\text{g}$ and $150 \pm 10 \mu\text{g}$) were loaded on Immobiline DryStrips pH 3 – 10 and pH 6 – 11, respectively.

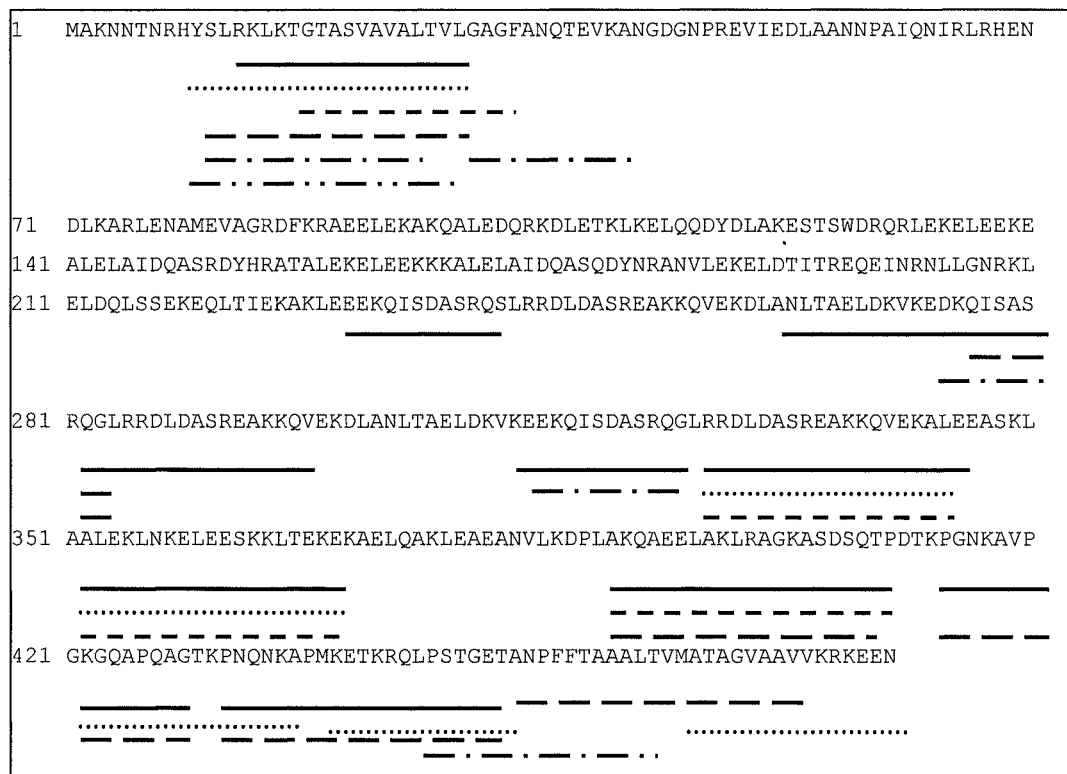


Figure 4.22 Identity of regions of highly- conserved regions in the M1 protein. Plasminogen-binding activity, ——— ; M-like proteins, - - - ; Fibrinogen-binding activity, ; ImmunoglobulinG binding activity, — — — ; ImmunoglobulinA binding activity, _ _ _ ; Fc- γ -receptor _ _ _ .



Figure 4.23 Identity of regions of highly- conserved regions in tRNA modification GTPase.

tRNA modification GTPase activity: — ; thiophene and furan oxidation protein: ··· , ThdF protein-tRNA modification: — · , GTP-binding protein engA: — · , tRNA synthase: — — , and Cytidylate kinase — · — .

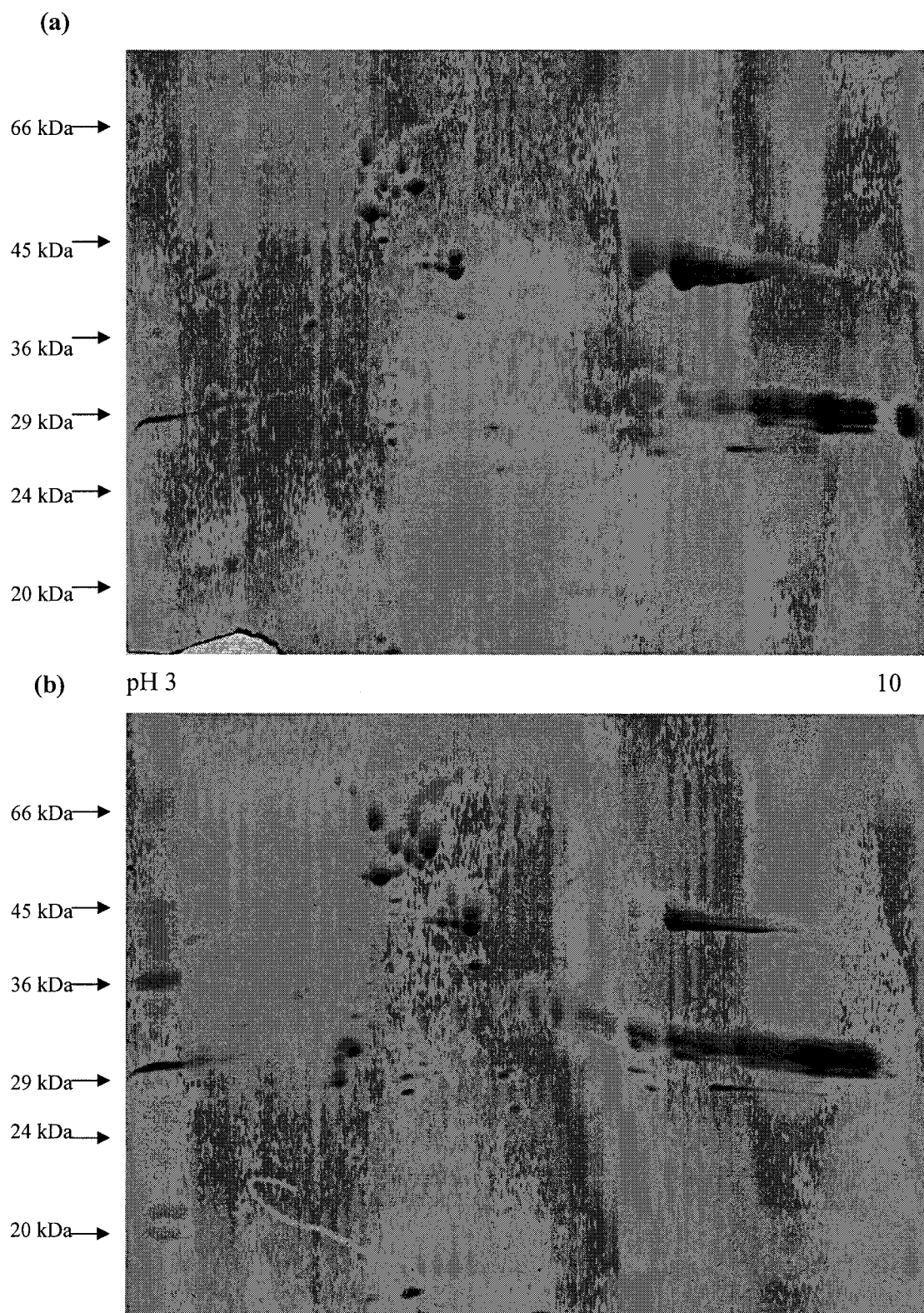


Figure 4.24 Comparison of secreted proteins from *S. pyogenes* SF370 culture supernatant in pH range 3 – 10.

(a) Secreted proteins from control culture supernatant; (b) Secreted proteins from HA-rich culture supernatant.

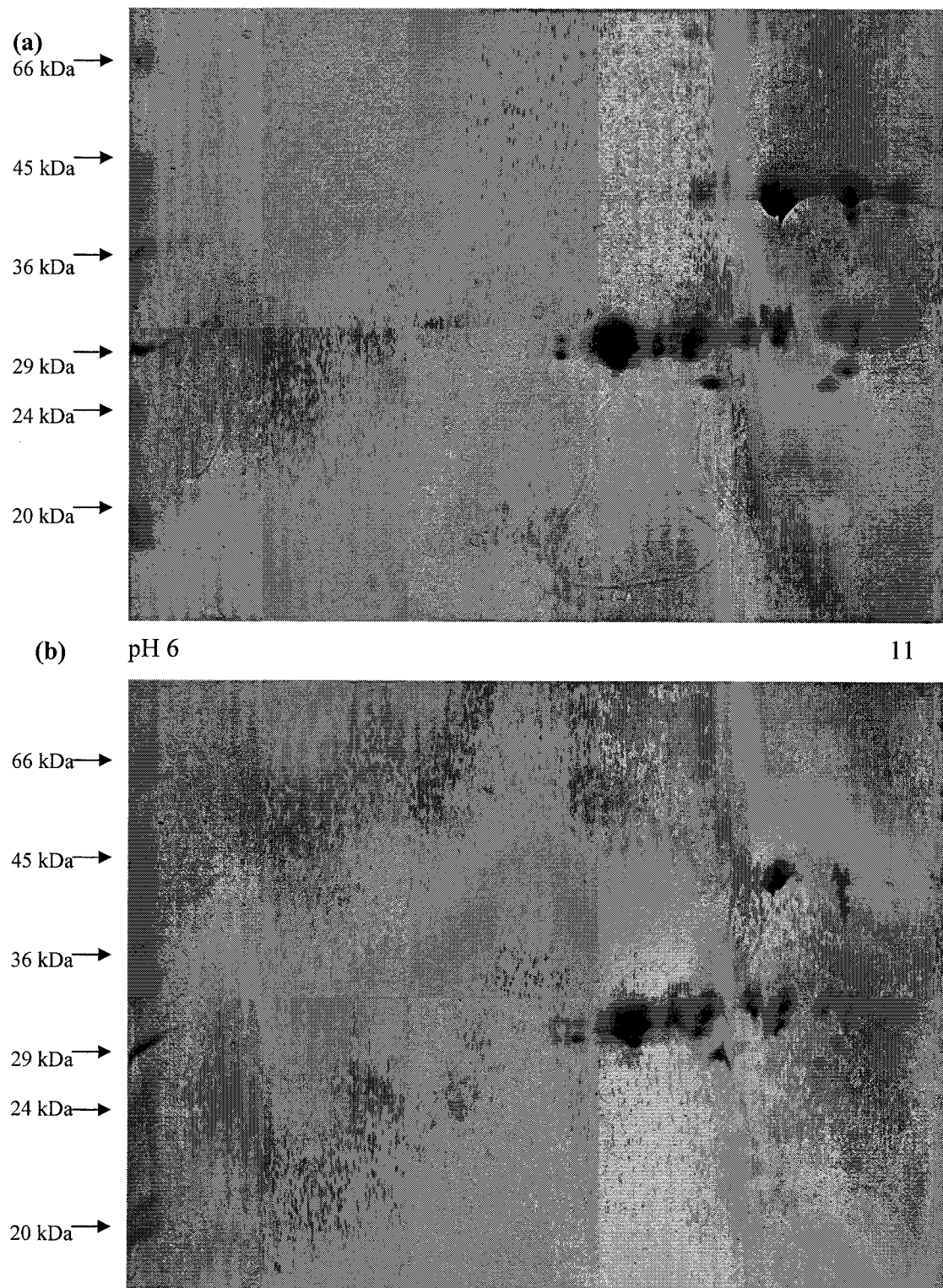


Figure 4.25 Comparison of secreted proteins from *S. pyogenes* SF370 culture supernatant in pH range 6 – 11.

(a) Secreted proteins from control culture supernatant; (b) Secreted proteins from HA-rich culture supernatant.

5 Discussions

5.1 Development of techniques for proteomics

5.1.1 Development of techniques for 2DE

2DE is considered to be a useful tool in proteomics for investigating biological process at the protein level. Although the concept of 2DE is simple, the techniques are variable from laboratory to laboratory, depending on the sample being investigated and the systems used.

Sample preparation is the first and most important step for successful 2DE, and it is critical for detailed visualization of expressed proteins on 2DE gels (Long et al., 2006). The general requirement for a sample is that it is soluble and denatured, which is achieved by the use of several agents, e.g. a chaotropic agent (such as urea) which is used to solubilise and denature proteins; a detergent (such as CHAPS) which is used to solubilise hydrophobic proteins and to minimize protein aggregation, a reducing agent (such as DTT) which is used to reduce disulphide bonds and allow complete protein unfolding, a buffer and ampholytes (i.e. IPG buffer) which are used to improve sample solubility and separation (Govorun and Archakov, 2002). However, sample preparation is very sensitive to slight variations in the concentrations of solubilising agents and therefore small differences can have a dramatic effect on the protein patterns on 2DE gels.

At the beginning of this project, several experiments were done to optimize the methods to prepare the proteins samples from different sources, i.e. cell-associated proteins from *E. coli* TOP10 and *S. pyogenes* SF 370, and secreted proteins from *S. pyogenes* SF 370. The concentration of 2 % (w/v) CHAPS in lysis solution and 0.2 % (w/v) DTT in rehydration solutions were used, for extraction of protein from *E. coli*, which is based on O'Farrell's lysis buffers (O'Farrell, 1975), and is the most popular sample solubilisation solutions currently used worldwide. Also, even though the protein sample was not treated with a 2-D Clean Up kit, lots of clean protein spots

were present on the gel image, and there was no obvious streaking phenomenon (data not shown). However, when the same solutions were used to extract the proteins from *S. pyogenes* SF 370, which is a Gram-positive bacterium and produces a large amount of extracellular polysaccharides, there were only a few spots visible on the 2DE gel (Figure 3.2). The loss of the protein spots was due to incomplete solubilisation of the proteins during extraction, and was improved by increasing the concentration of CHAPS to 4 % (w/v) in the lysis solution (Figure 3.4), and increasing the concentration of DTT in rehydration solutions, which helps to fully reduce disulphide bonds between proteins and therefore resulted in clean protein spots on the 2DE gels (Figure 3.6). Also, a 2-D clean up kit was used to remove the contaminations from the samples which reduced the streaking phenomenon. For isolation of secreted proteins from the supernatant of a *S. pyogenes* SF370 culture, TCA/acetone precipitation was used. The above experiments proved that there is no single method that worked for all protein extractions, and slight variations in the concentrations of the solubilising agents can make a big difference in the quality of the 2DE gels.

In addition to working out an appropriate method for each organism, reproducibility is another critical issue in sample preparation for 2DE. However, slight technical variability in the process of sample preparation results in large changes on the 2DE gels (Choe and Lee, 2003). Especially when comparing differentially expressed proteins from one organism. If the proteins from the different sources, for example HA-rich media and control media in this study, are extracted at different time or under different conditions, the changes on the 2DE gels may be caused by the variation during sample preparation rather than 'real differences', i.e. differential expression due to the present of HA in the media. Therefore, in this project, for each set of comparison, samples were prepared at the same time and under the same conditions. In addition, eight sets of comparison were performed to confirm the changes were 'real'.

To increase the reproducibility of gels, similar amount of proteins were loaded on the gels, which was achieved by measuring protein concentration in the sample using the Bradford assay or the 2-D Quant kit prior to loading. The Bradford assay is the most popular method for measuring the concentration of proteins in the solution. However, the protein samples for 2DE normally contains urea, CHAPS and DTT, which

interferes with Coomassie-based protein assays, reducing colour development and causing precipitation of the assay reagent (Ramagli and Rodriguez, 1985). Therefore, the 2-D Quant kit was employed, which is compatible with up to 8 M, 4 % (w/v) CHAPS and 1 % (w/v) DTT.

The Multiphor II system and Protean II XL cells were used in the project for 1st dimension – IEF, and 2nd dimension – SDS-PAGE respectively. IEF was performed at different pH ranges. The wider pH range of 3 -10 offered an overview of the state of the proteome. However, due to the insufficient spatial resolution, wide range IEF only revealed a small percentage of the whole genome (Figure 3.6). Therefore, the proteome map was ‘zoomed in’ to pH range 4-7 and 6-11.

There are two different methods that can be used for IEF sample loading, rehydration loading (Rabilloud et al., 1994) and cup loading (Gorg et al., 1988). The choice depends on the pH range of the Immobiline DryStrip gels to be used. For wide range, e.g. pH range 3 – 10, both sample loading methods can be used. However, several studies have reported that rehydration loading is preferred, because this method can be used to load more proteins, up to 200 µg, and this method has also been shown to improve the resolution of 2DE (Rabilloud et al., 1994, Gorg et al., 2000), which is consistent with the results I obtained from experiment No.1 and No.2 (data not shown). Also for pH range 4 -7, rehydration loading produced good quality gels (Figure 3.8 and Figure 3.9); however, when the IEF was performed on basic pH gradient such as pH 6-11, cup loading was preferred for all the samples (Gorg et al., 2000).

Two different staining methods, Coomassie blue staining (Figure 3.4) and colloidal Coomassie blue staining (Figure 3.5), were used in this study to visualize the 2DE gels. Both these staining methods are compatible with MS analysis and, as the more sensitive the method the better the possibility of visualizing low-abundance proteins in a sample, colloidal Coomassie blue staining was used to visualize the gels. Colloidal Coomassie blue staining is a modified Coomassie blue staining method, i.e. Coomassie Blue G 250 instead of Coomassie Blue R250 is used, and 20 % (v/v) methanol is added, increasing the sensitivity of the staining down to the nanogram level.

5.1.2 Analysis of 2DE gels using PDQuest™ software

A big challenge for analysis of 2DE gels is that gels suffer from high variation since the intensity of spots, which represent the expression level of the protein, can vary significantly on different gels (Houtman et al., 2003). As mentioned in section 4.1.1, this variation may be introduced during sample preparation and protein loading and also may be caused by staining. Therefore, in this study a large number of replicate gels from the same conditions were produced to obtain statistically significant changes in protein levels between different groups, i.e., sample from control media versus sample from HA-rich media. Eight sets, totalling 16 2DE gels, of comparisons were performed, e.g., eight gels of sample from control media and eight gels of samples from HA-rich media. This compares very favourably to other studies where only three replicates were performed (Nakamura et al., 2004, Cash et al., 1999).

Although 2DE was optimized in this project, in order to minimize artificial variation, there was still the requirement to normalize 2DE gels after the protein patterns were compared and matched. The proteins were then analyzed for changes in levels of expression using the Wilcoxon two-sample test, which is a nonparametric statistic employed ideally for data sets of less than 20 and greater than 5 (Houtman et al., 2003). The Wilcoxon two-sample test is a more critical statistical method for identifying differentially expressed proteins. Although fewer differentially expressed proteins (20 proteins *cf* 40) were identified using this statistical method compared to the “2-fold changes” method that is frequently used in proteomic analysis, the significance of this differential expression is much higher.

5.1.3 Development of techniques for static nanospray LCQ Advantage MS

As mentioned in the Introduction section, the development of ESI makes ESI MS a useful tool in the analysis of proteins, and this application was first demonstrated in the late 1980s by Fenn et al. (Fenn et al., 1989). In 1996, low flow rate electrospray was developed. Using this system, the flow of the spray can be reduced to nanoliters per minute and small sizes of droplets can be generated which leads to extension of the

time for mass spectrometric analysis. Therefore low abundance proteins from limited amount of sample can be analysed. For this reason, nanospray MS has become a widely used technique for many proteomics studies.

Static electrospray is an off-line system. The sample is loaded directly into the static emitter needle and the flow is achieved passively rather than using solvent pumps. The static nanospray ionization source has been proved to be a very stable source and can spray a wide variety of buffers in positive and negative modes (Wilm and Mann, 1996). This system compares well to the nano-LC/MS system, which is widely used in proteomics studies, and has several advantages including: (a) complete absence of any cross contamination because the static emitter needle used is not recycled; (b) few opportunities for sample loss, (c) more time for ion trap spectrometric analysis for each scan as flow rate is lower than any nano-LC/MS system, (d) less overall time for analysis of each sample than nano-LC/MS, and (e) it is a cheaper approach for identifying proteins.

However, low salt tolerance is a major limitation of nanospray ionization. Normally, when samples contain high salt concentrations ($> \sim 1$ mM), analyte ion formation is effected (Wilm and Mann, 1996). To investigate the salt tolerance of my static nanospray system, serial dilutions of protein digests containing varying concentrations of NaCl were analysed (Figure 3.17 and Figure D1-D8). Peptide peaks were visible when the solvent contained less than 1 mM NaCl, but they started to disappear when the NaCl increased to ~ 10 mM, and they could not be detected at all when the concentration was ~ 100 mM. These results show that the static nanospray LCQ Advantage MS has a similar salt tolerance to LC nanospray mass spectrometry.

In-gel digests from 2DE gels contain high concentrations of not only salt but also detergents. Therefore the detergent tolerance of the system was also tested (Figure 3.17 and Figure D1-D8). When the solvent contained less than 0.1 mM SDS, the peptide peaks were readily identified, but they started to disappear when the SDS concentration increased to ~ 1 mM. However, when the SDS concentration was ~ 10 mM, the peptide peaks could not be detected at all. Therefore, the treatment of digests with ZipTips to remove the contaminants is a necessary step prior to MS analysis.

As can be seen from the results shown in Figure 3.17, the higher concentration of analytes the better MS profile produced. For the purpose of protein identification from 2DE gels, the digests of individual protein spots from 6 of the 8 replicate gels were combined, and submitted on static nanospray LCQ Advantage MS.

5.1.4 Protein identification

With the MS system used in this project, proteins were identified from peptide information obtained by MS/MS. Peptide information was acquired by selection, isolation and fragmentation of peptides (precursor ion) within the mass spectrometer. The fragmentation occurs mainly in the peptide bond of the amino acid chain backbone, and the fragmented patterns are reproducible and, in general, predictable (Eng et al., 1994). Computer programs have been developed to calculate the predicted fragment ions for a given peptide to facilitate database matching with experimental MS/MS data. This process is also called peptide fragmentation fingerprinting (Chamrad et al., 2004).

TurboSEQUENT ranks its results according to a peptide probability value that the experimentally measured mass spectra correspond to virtual mass spectra which have been obtained by a virtual digest of the protein database. The major challenge of MS/MS based protein identification is that peptide masses determined by MS are generally not unique which means each measured mass can randomly match a peptide from the sequence database. The project used a streptococci database rather than a *S. pyogenes* SF370 database and allowed missed cleavage sites, increased the number of random matches, and reduced the discrimination as well as the significance of the search results. As a result, there remains a certain risk of obtaining a false positive (Chamrad et al., 2004). Therefore certain criteria were used to distinguish correct identifications from false positives, e.g., proteins must be from the SF370 strain of the M1 serotype of GAS; and peptides must have a $X_{corr} > 1.5$ for singly charged peptides, > 2.0 for doubly charged peptides, and > 2.5 for triply charged peptides, with at least 2 peptides successfully matched; and/or must have a $\Delta C_n > 0.2$ with at least 2 peptides successfully matched (section 2.10.1.3). When the peptides satisfied the X_{corr} and/or ΔC_n criteria and more than one peptide was identified for each protein identified there was a higher confidence in the results.

5.1.5 Functional analysis of proteins

In the last 10 years, the major bioinformatics application of proteomics has been protein identification by searching MS/MS data against genome or protein database. As there has been an accumulation of molecular biology information, there is now a critical need for bioinformatic tools that can combine biochemical data about an unidentified peptide from different experiments and search databases to identify the potential functions of the proteins in biology system (Figure 5.1) (Kanehisa and Bork, 2003, Dutt and Lee, 2000).

The similarity search against primary sequence database is used for functional prediction of a peptide or protein as long as the database is well annotated, although it is increasingly more difficult to maintain an up-to-date sequence database owing to the varying quality and ever-increasing amounts of sequence data. Therefore, recently it is the trend to rely more on searches against secondary databases, which contain conserved protein domains and functional identities (Kanehisa and Bork, 2003).

TSSview™ v1.0 is software tool for identifying regions of evolutionary conservation in proteins as determined by high-sensitivity data-mining conducted on approximately three million UniProt entries and their corresponding coding and non-coding regions in each of the six reading frames. By using TSSview™ v1.0 software, the highly-conserved regions of protein sequences can be identified. This functional analysis of proteins plays an important role in proteomics analysis, as far as understanding the biology is concerned.

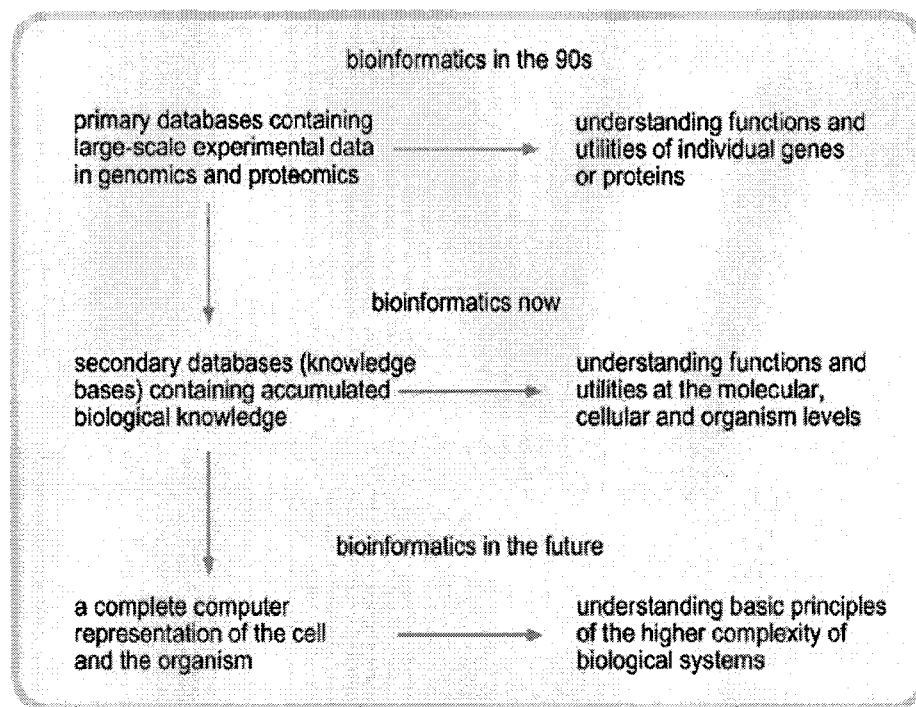


Figure 5.1 A view of bioinformatics past, present and future (Kanehisa and Bork, 2003)

5.2 Proteomics analysis of *S.pyogenes* SF370 growing in HA-rich media

S. pyogenes SF370 is a M1 serotype GAS, which is among the most common and versatile of all extracellular bacterial pathogen, and it is responsible for a wide spectrum of human diseases, ranging from trivial to lethal (Bisno et al., 2003). It primarily causes localised infections of the mucous membranes, tonsils or the skin, but occasionally invades deeper sites, causing impetigo/pyoderma, cellulitis, necrotising fasciitis, scarlet fever, septicaemia, pneumonia and meningitis (Bisno et al., 2003, Cunningham, 2000). A wide variety of proteins and toxins produced by this pathogen are known or suspected to contribute to the virulence of GAS (Figure 1.4). Since GAS infections cause half a million deaths every year (Carapetis et al., 2005), it is a worldwide problem and continuing research is needed to keep pace with this evolving pathogen.

In addition to the classic molecular biological approach, such as generation and characterization of isogenic mutants, the expression of gene products involved in virulence have been studied using transcriptomics and proteomics. Smooth and his group investigated global differential gene expression in response to growth temperature alteration (Smoot et al., 2001, Cash et al., 1999). Voyich et al. reported that the GAS genes are differentially regulated during phagocytic interaction with human PMNs comprise a global pathogen-protective response to innate immunity (Voyich et al., 2003). Sumby et al. studied the difference of total gene expression between different GAS strains isolated from pharyngeal disease using bacteria in soft-tissue infection models of disease (Sumby et al., 2006). Lei et al. using proteomics technology compared the culture supernatants of wild-type GAS strains and isogenic mutant strains and identified previously undescribed extracellular proteins, including several that are immunogenic in the course of host-pathogen interactions (Lei et al., 2000). Cole and his group identified 74 distinct cell wall-associated proteins, and 66 of them were novel (Cole et al., 2005). Chaussee et al used a proteomic approach to compare proteome map of GAS wild-type strain and an *rgg* mutant strain and concluded that Rgg is a global regulatory factor that contributes to growth phase-dependent synthesis of proteins associated with secondary metabolism and oxidative and thermal stress responses (Chaussee et al., 2004). In addition, proteome analysis has been used to confirm findings from transcriptomics experiments and to distinguish

between transcriptional and post-transcriptional regulatory events (Graham et al., 2001).

To understand the molecular mechanism of GAS virulence, studies focusing on how GAS reacts to its environment and the mammalian host are required. During GAS infection and invasion of human tissues, the expression of gene products of GAS are altered, including virulence factors and other proteins involved in metabolism. As HA is a major component of human connective tissue, the destruction of connective tissue and the HA therein, is a key element of GAS pathogenesis. Therefore, to examine simultaneous changes in expression of multiple virulence factors, GAS was propagated in HA-rich media, in which HA was added to growth media in an attempt to create a simple biological system that could reflect some elements of GAS pathogenesis, and proteomic analysis was performed.

5.2.1 Growth of *S. pyogenes* SF370 in HA-rich media and control media

The growth of *S. pyogenes* SF370 in both HA-rich media and control media was investigated (Figure 4.1). There was no obvious difference between the growths of GAS in these media, even though HA is a high-molecular-mass polysaccharide, and could be a carbon/ energy source for the growth of GAS.

To determine whether *S. pyogenes* uses HA as a carbon/energy source, the consumption of carbohydrates in HA-rich media was investigated. Because HA-rich media is a non-defined media (i.e. BHI with added HA), it was not surprising to find that the concentration of uronic acid was 58 mg/ml in the control media (BHI with no added HA). The consumption of carbohydrates containing uronic acid in HA-rich media by *S. pyogenes* was more than in control media; eight mg in every ml of HA-rich culture, and four mg in every ml of control culture. This indicated that some carbohydrate containing uronic acid moieties, which might be HA, was consumed during the growth of *S. pyogenes* in HA-rich media. However, the investigation of consumption of total reducing sugar, rather than just uronic acid containing molecules, in the same media show that there were no obvious difference in the consumption of total reducing sugar between HA-rich media and control, which is consistent with the

results in Figure 4.1. It is known that bacteria are capable of using numerous carbohydrates as sources of carbon and energy. Several studies have shown that during the growth of the Gram-positive bacteria, preferred carbohydrates are utilised when a wide variety of carbon/energy sources are present in the growing environment. This ability depends on carbon catabolite repression (Iyer et al., 2005, Stulke and Hillen, 2000). In HA-rich media HA may be one of many carbon sources, and compared to other carbohydrates in the control media, may be a preferred carbon source for *S. pyogenes* growth. However, because there was no significant difference in the consumption of total reducing sugar between HA-rich media and control media, differential expression of proteins is therefore likely to be the result of bacterial “reaction” to HA, similar to the GAS infection, rather than the result of increased growth due to the presence of HA.

5.2.2 Analysis of expression of hyaluronate lyase by *S. pyogenes* SF370

Hyaluronate lyase is an extracellular factor present in the culture supernatants of GAS (Hynes and Walton, 2000), and its substrate is HA, which as well as being the major component of human connective tissue, is also found as the capsule of GAS (Hynes et al., 1995). *S. pyogenes* SF370 encode four hyaluronate lyases, HylP1, HylP2, HylP3 and HylA. HylP1 has been demonstrated as a hyaluronate lyase, and it is produced by prophage of this organism. This enzyme has been structurally and functionally characterised, and it is suggested that the role of HylP1 is not HA degradation *per se* but reduction of the viscosity of HA to allow phage penetration of the *S. pyogenes* cell wall (Smith et al., 2005).

Several experiments have been done in this project to determine whether *S. pyogenes* SF370 produced any hyaluronate lyase activity. The results of the DNSA reducing sugar assay and the 232 nm assay show that neither the *S. pyogenes* SF370 cultures grown in control media nor the *S. pyogenes* SF370 cultures grown in HA-rich media, produced any such enzyme activity (section 4.2.3 and Figure 4.2). Also the 2DE gels of both supernatants of *S. pyogenes* SF370 in pH range 6 – 11 (Figure 4.25) show that there were no protein spots in the appropriate area of the gel (i.e. pH 9.0 – 9.3, and 36,000 – 40,000 kDa). However, according to studies of Starr and Engleberg, GAS can

only use HA as a carbon/energy source when hyaluronate lyases are expressed (Starr and Engleberg, 2006). Since I have shown that HA was consumed in cultures of *S. pyogenes* SF370 grown in HA-rich culture, it was assumed that hyaluronate lyases were produced in the HA-rich media supernatant. However, it has been reported that during the stationary phase of growth of GAS, reduced levels of several virulence factors including hyaluronate lyases are expressed (Leonard et al., 1998), and since all the samples I prepared were at stationary phase, it is possible that the level of hyaluronate lyases in my system was not high enough to be detected. This is also consistent with the result of Starr and Engleberg, who showed that only 10 % of clinical isolates produce detectable levels of hyaluronate lyase activity (Starr and Engleberg, 2006). However, it would be worth measuring the level of hyaluronate lyases produced in the exponential phase in the future.

5.2.3 Analysis of differentially expressed proteins in HA-rich media

To examine the response of GAS to the presence of HA in the media, and to determine the differentially expressed cell-associated proteins, 2DE was performed. The proteome map in pH range 3-10 was performed first, and greater than 90 % of the protein spots were shown in the acid area (Figure 3.6). Therefore the comparison was performed on 2DE gels in pH range 4 – 7. According to Wilcoxon two-sample test, 18 proteins spots were shown differentially expressed by *S. pyogenes* SF370, and 12 proteins were up regulated in HA-rich media, 6 proteins were down regulated in the HA-rich media. Among the 18 protein spots, 13 proteins were identified. 11 proteins were identified according to predefined Xcorr and/or ΔCn values, and another two proteins were added to the list of identified proteins because TSSview™ showed that the single peptide identified from these two proteins using MS is a functionally highly-conserved peptide.

The observed up-regulated proteins in HA-rich media are involved in (a) protein and DNA synthesis, i.e. 50 S ribosomal protein L10 (Figure 4.12), the tRNA modification GTPase (Figure 4.10), a putative formate-tetrahydrofolate ligase (Table 4.3) and the carbamoylphosphate synthase small subunit (Figure 4.15), (b) transport, i.e. the phosphohistidine-containing protein, HPr (Figure 4.5), a component of the

phosphotransferase system (LiCalsi et al., 1991) and (c) glycolysis, i.e. glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Figure 4.6). Although there was no statistical difference between the growths in the two different media, GAS grown in HA-rich media reached higher optical densities compared to that grown in control media. In addition, a study by Chaussee et al. showed that even though growth rates of wild type and *rgg*-negative mutants of GAS are similar, they produce different amounts of biomass (Chaussee et al., 2004). It was therefore not surprising that the up-regulation of proteins involved in transport and biomass/energy production was observed. These observations are in line with those of Voyich et al. (Voyich et al., 2003) who also observed the up-regulation of many of these proteins following analysis of a phagocytic interaction of GAS with human PMNs. In addition, it has been reported that HPr is involved in the mediation of carbon catabolite repression when Gram-positive bacteria grow in media with multiple carbon sources (Warner and Lolkema, 2003, Titgemeyer and Hillen, 2002, Galinier et al., 1997). Also the up-regulation of proteins involved in carbohydrate metabolism was observed in several transcriptome studies of GAS infection in an animal infection model and human blood model (Virtaneva et al., 2005, Graham et al., 2005).

Two LPXTG motif-containing cell-wall anchored proteins were shown to be up-regulated in HA-rich media: M1 protein precursor (Figure 4.7) and a collagen-like surface protein (Figure 4.14). LPXTG motif-containing cell-wall anchored proteins have been shown to play important roles in pathogenesis, such as resistance to phagocytosis, adherence to plasma and extracellular matrix proteins, and degradation of host proteins (Hynes, 2004). As LPXTG motif-containing proteins are covalently attached to the cell wall, I was surprised to see these proteins in the cell-associated protein fraction, as I did not expect them to be released from the cell wall during protein preparation. However, the sonication step used during protein preparation may have released these proteins from the cell wall, as seen elsewhere by Dryla et al (Dryla et al., 2005). Alternatively, it is possible that the precursor forms of these proteins were derived from the cytoplasm and not the cell wall.

The first LPXTG motif-containing protein identified as up regulated protein in HA-rich media was M1 protein, the major GAS virulence factor (Cunningham, 2000). M1 proteins play several roles in GAS pathogenesis; they protect the bacteria from

phagocytic clearance by interfering with complement-mediated uptake (Husmann et al., 1997) and, during throat infection, they promote bacteria-bacteria interaction following attachment to tonsillar epithelial cells (Caparon et al., 1991). Functional analysis by TSSview™ of the sequence of the M1 protein showed that this protein has several conserved regions (Figure 4.22). As shown in this figure, conserved areas of M-like proteins were discerned. M-like proteins share a similar structure with M proteins and are encoded by *emmL* genes. Both M protein and M-like proteins have been shown to have the ability to bind to the Fc region of immunoglobulin (Guyre et al., 1990, La Penta et al., 1994, Sjobring et al., 2002), and this ability can be further subdivided into those proteins that bind Immunoglobulin A or proteins that bind to Immunoglobulin G (Frithz et al., 1989, Gomi et al., 1990, Burova et al., 2005, Podbielski et al., 1996). The ability of M1 proteins to bind immunoglobulin plays an important role in anti-phagocytosis. The conserved regions that facilitate immunoglobulin binding were shown to be present in the GAS M1 protein sequence. In addition, the M1 protein sequence was shown to have several conserved regions for plasminogen-binding function. This function has been shown to facilitate the accumulation of plasmin on the GAS surface, and can therefore help break host barriers and promote migration of GAS, which is important for the invasive GAS infections (Berge and Sjobring, 1993). The M1 protein also possessed conserved regions for fibrinogen-binding activity. This activity can precipitate fibrinogen that prevents the killing of GAS in host blood (Johansson et al., 2004). The observation of up-regulated M1 protein is consistent with the results of transcriptome studies by Virtaneva et al. and Graham et al.. Virtaneva et al. analysed the transcriptome of GAS during an 86-day infection protocol in 20 cynomolgus macaques with experimental pharyngitis. They found expression of M protein increased during the infection and reached the highest level in acute phase of disease (Virtaneva et al., 2005). Graham et al. observed the up-regulated expression of M protein during GAS wild type growth in human blood (Graham et al., 2005).

The up-regulation of another LPXTG motif- containing cell-wall anchored protein, i.e. a collagen-like surface protein was also observed in HA-rich medium. Lukomski et al. demonstrated that a collagen-like surface protein mutant possessed a significantly reduced ability to adhere to human epithelial cells grown in culture and was significantly less pathogenic in a mouse model of soft tissue infection, which suggested that collagen-like surface protein is a GAS virulence factor that participates

in the adherence of GAS to host cells and soft tissue pathology (Lukomski et al., 2000). However, in the same paper, they also suggested that no *scl* transcription was detected in stationary phase cultures. In my system, the observation of this protein may be due to its accumulation from the early growth phase.

Interestingly, the glycolytic enzyme, GAPDH (Wilkins et al., 2002, Pancholi and Fischetti, 1992) and the transport protein, HPr (Dixon et al., 2001) have been identified at the cell surface of pathogenic streptococci, with GAPDH shown to be implicated in pathogenesis. It has been demonstrated that GAPDH was able to bind to fibronectin, lysozyme and cytoskeletal proteins and this multiple binding capacity of the GAPDH has been postulated to play a role in the colonization, internalization and the subsequent proliferation of GAS (Pancholi and Fischetti, 1993). Additionally, GAPDH produced by GAS has been found to have adenosine diphosphate-ribosylating activity (Pancholi and Fischetti, 1993). Since adenosine diphosphate-ribosylation is an important component of intracellular signalling events, GAPDH has been shown to have a role in mediating cell-to-cell communication between GAS and pharyngeal cells (Pancholi and Fischetti, 1997). Recently, it has been found that GAPDH helped GAS to bind to C5a and subsequently results in which is degraded by C5a peptidase (Terao et al., 2006)

Interestingly, two other proteins that were up-regulated in HA-rich media have been shown to be involved in signal transduction. In addition to the transport protein HPr, which was found to have a role in signal transduction that involves the regulation of carbon catabolite repression, the histidine protein kinase (Figure 4.9), which was up-regulated in HA-rich media, plays a role in signal transduction. In general, histidine protein kinase functions as an environment sensor of two-component signal transduction regulatory systems, which are used by bacteria to sense and respond to specific environmental changes (Wang et al., 2001). In two-component signal transduction regulatory systems, the membrane-located histidine kinase sensor senses the specific environmental signal, triggers phosphorylation of a DNA-binding response regulator and allows the regulator to bind to DNA. Such binding activates or represses the nearby promoter (Dalton and Scott, 2004). A two-component regulatory system, designated *covRS* (*cov*, control of virulence; also known as *csrRS*) in GAS is well characterized. CovS is a membrane-located sensor and responds to environmental changes and CovR functions as a streptococcal global regulator. Expression microarray

analysis found that this system, *in vitro*, influences (directly or indirectly) transcription of 15% of all GAS genes, including many that encode surface and secreted proteins mediating host-pathogen interactions (Graham et al., 2001). For example, it negatively controls the expression of the HA capsule, Spe B, streptokinase, SLS, and streptodornase (Musser and DeLeo, 2005, Senadheera et al., 2005). The increased expression of CovS was also observed in transcriptome study by Virtaneva et al. during the infection of cynomolgus macaques, especially in the acute phase of the disease (Virtaneva et al., 2005).

Several proteins were also identified as down-regulated. A regulator of a streptococcal virulence factor, RopA (Figure 4.13), was down-regulated in HA-rich media. RopA is a regulator of streptococcal pyrogenic exotoxin B (SpeB) and assists SpeB in translocation via the secretory pathway and functions as a molecular chaperone to achieve an active conformation of SpeB (Collin and Olsen, 2003). The down regulation of RopA may suggest that the level of expression of SpeB may have decreased in my system. However, SpeB was not identified in my studies, as it is secreted and only cell-associated proteins have been analysed. Therefore, it would be worth analysing the differentially expressed secreted proteins in the future.

The down regulation of a transcriptional repressor, CodY (Figure 4.16), was as observed in my system. Malke et al. have shown that CodY is a growth phase-dependent positive transcriptional regulator of *mga* and *CovR* (Malke et al., 2006), which positively affect the expression of six virulence factors (Hyaluronate Synthase A, Immunoglobulin G Degrading Endopeptidase, Pyrogenic Exotoxin H, C5a peptidase, a cell surface proteinase, and a collagen-like surface protein). Therefore the down regulation of CodY might suggest that these virulence factors should be down regulated. But in my system I have not identified the down regulation of these preteins, and except for the up regulation of a collagen-like surface protein. This, therefore seems to be in conflict with the down regulation of CodY. However, the regulation of *mga* and *CovR* by CodY has been shown to occur at the mid-exponential phase of cell growth (Malke et al., 2006), while my cells were harvested at early stationary phase. Interestingly, at early stationary phase, CodY has also been shown to act as a guanosine 5-triphosphate-binding protein (Ratnayake-Lecamwasam et al., 2001) that senses the intracellular guanosine 5-triphosphate concentration as an

indicator of nutritional limitations. It is therefore possible that the observed down-regulation of CodY is due to a reduction in the levels of guanosine 5-triphosphate in my system at early stationary phase. The down-regulation of CodY was also observed by Graham et al during wild type GAS was exposed to human blood (Graham et al., 2005).

TurboSEQUEST™ identified two hypothetical proteins, the functions of which have not been annotated in the genome database. By analysis of these proteins using TSSview™, it was found that hypothetical protein SPy1143 (Figure 4.8) has conserved regions that are also present in SUA5 protein. This protein is required for the normal growth of yeast cells (Na et al., 1992), and similar sequence have been found in the M28 GAS serotype (Green et al., 2005). However the function of SUA5 is not yet known, therefore whether this protein is involved in the GAS pathogenesis still needs to be elucidated.

Another hypothetical protein, SPy1262 (Figure 4.11) has been shown by TSSview™ to include some of the conserved regions of the general stress protein, Gls24. This protein, in *Enterococcus faecalis*, has been found to be implicated in virulence as well as stress response (Teng et al., 2005, Nannini et al., 2005) and may therefore have a similar role in GAS. The observed up-regulation of this protein is consistent with the result of the transcriptome study performed by Graham et al. They found that SPy1262 was up-regulated in the late stage of GAS exposure to human blood (Graham, et al., 2005).

5.3 Conclusion

Proteomics, as a modern molecular biological technology that has been widely used in biology systems studies, especially in the investigation of mechanisms of bacterial pathogenesis. Streptococci are Gram-positive bacteria, and are responsible for a wide range of human diseases and animal disease. These diseases cause millions of deaths and huge economical loss. In the last few decades, large amounts of work have been done to investigate the pathogenesis of these bacteria. However, the use of classical molecular biological approaches alone is insufficient in furthering our knowledge of the mechanisms of pathogenesis. As the number of genome sequences increase,

proteomics technology is starting to be used in this area of pathogenesis analysis over the last few years.

To apply proteomics technology in the studies of streptococcal pathogenesis, several techniques were developed in this project, including separation of protein from different sources using 2DE, 2DE gel analysis using PDQuest software, in-gel trypsin digestion, protein identification using static nanospray LCQ Advantage MS, database searching using TurboSEQUENT and functional identification and analysis of differentially expressed protein using TSSviewTM software. From the results of this project, it has been proved that the coupling of 2DE with static nanospray MS methodologies is a powerful tool in high-throughput proteomics.

Proteome maps of difference pH ranges for cell-associated proteins from *S. pyogenes*, and for proteins from culture supernatants of *S. pyogenes* were successfully constructed. Also the comparison of differentially expression proteins was performed for *S. pyogenes* SF370 grown in control media and HA-rich media, which was used to create a simple biological system that reflected some elements of GAS pathogenesis. The proteins were identified and functionally analysed. Some of these proteins are major virulence factors of GAS, such as the M1 protein and a collagen-like surface protein, some are regulators of virulence factors, such as RopA, and some are proteins which are not commonly associated with virulence, e.g. the glycolytic enzyme GAPDH, but have been shown to play important roles in streptococcal pathogenesis. Interestingly, I also identified a hypothetical protein, which may be involved in GAS pathogenesis.

Unlike many models of streptococcal pathogenicity, I have shown that growing GAS in HA-rich media is readily amenable to proteomic analysis due to the relatively large amounts of biomass propagated. The fact that I have observed the up-regulation of several recognized virulence factors indicates that this simple biological system reflects several elements of GAS pathogenesis in humans and is therefore a valuable tool in the identification of potential virulence factors.

5.4 Future work

As the developments in proteomics technology continue, numerous opportunities for future studies exist in the areas of research covered in this thesis.

On a technical side, more sensitive staining methods could be applied for the visualization of 2DE gels. For example fluorescent dyes, such as SYPRO Ruby have been used in some proteomics studies (Renzone et al., 2005), which is a more sensitive staining method and more compatible with MS analysis of proteins (Berggren et al., 2000).

To address the problem caused by the gel-to-gel variability the technique called two dimensional difference gel electrophoresis (2D-DIGE) could be applied in the future, since it is based on the use of a single gel for the simultaneous separation of multiple protein samples, followed by the independent visualization of each individual sample (Renzone et al., 2005). 2D-DIGE can be achieved by labeling the control and comparison samples with different dye reagents, such as Cy3 and Cy5 respectively prior to IEF and analyzed simultaneously by a single 2DE gel (Yan et al., 2002, Alban et al., 2003). An increase of confidence in real biological variations could be achieved by using a single gel for cross sample comparison.

As a large amount of proteins will be processed in the future, nano-LC/MS is a necessary tool for protein identification. The nano-flow rate LC system (such as the Ultimate 3000 system from Dionex) can be linked to the LCQ Advantage MS, and large batches of samples can be run overnight, or over the weekend using the autosampler.

Also the multi-dimensional peptide identification technology (MudPIT) which combines strong cation exchange and reverse phase separations in an online LC/MS experiment could be applied in my proteomics study. The use of two-dimensional peptide separations results in greater chromatographic peak capacities and protein identifications, and also requires more analysis time and results in greater numbers of MS/MS spectra. The digests of protein mixture could be applied to LC/LC-MS/MS

system without separation using 2DE, which also increase confidence in real biological variations by avoiding the artificial difference caused by gel-to-gel variability.

On an experimental design side, for *S. pyogenes* SF370, the identification and functional analysis of the differentially expressed secreted proteins that appeared on proteome maps of culture supernatant in control media and HA-rich media (Figure 4.24 and 4.25), the studies of the expression of cell-associated proteins and secreted proteins at the different growth phases, i.e. exponential and stationary phases, would provide more information. As most GAS virulence factors are expressed during exponential phase of growth (Cunningham, 2000), the analysis of differentially expressed proteins due to the addition of HA at that stage are more likely to determine more elements of GAS pathogenesis. Also membrane proteins are often involved in pathogenesis; therefore development of non-IEF techniques to study membrane proteins, as membrane proteins can not be separated using IEF, would be of interest. The potential functions of hypothetical proteins assigned using TSSview software could be examined using classic molecular approaches or proteomics studies of the mutants.

6 References

- ALBAN, A., DAVID, S. O., BJORKESTEN, L., ANDERSSON, C., SLOGE, E., LEWIS, S. & CURRIE, I. (2003) A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics*, 3, 36-44.
- ALBERTI, S., ASHBAUGH, C. D. & WESSELS, M. R. (1998) Structure of the has operon promoter and regulation of hyaluronic acid capsule expression in group A Streptococcus. *Mol Microbiol*, 28, 343-53.
- ALOUF, J. E. (1980) Streptococcal toxins (streptolysin O, streptolysin S, erythrogenic toxin). *Pharmacol Ther*, 11, 661-717.
- AMERSHAMBIOSCIENCES (Ed.) (2001) *2D Electrophoresis using immobilized pH gradient: Principles and Methods*.
- ANDERSON, N. G. & ANDERSON, N. L. (1996) Twenty years of two-dimensional electrophoresis: past, present and future. *Electrophoresis*, 17, 443-53.
- ANGELICHIO, M. J. & CAMILLI, A. (2002) In vivo expression technology. *Infect Immun*, 70, 6518-23.
- ASHBAUGH, C. D., WARREN, H. B., CAREY, V. J. & WESSELS, M. R. (1998) Molecular analysis of the role of the group A streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. *J Clin Invest*, 102, 550-60.
- AZIZ, R. K., PABST, M. J., JENG, A., KANSAL, R., LOW, D. E., NIZET, V. & KOTB, M. (2004) Invasive M1T1 group A Streptococcus undergoes a phase-shift in vivo to prevent proteolytic degradation of multiple virulence factors by SpeB. *Molecular Microbiology*, 51, 123-134.

- BEACHEY, E. H. & OFEK, I. (1976) Epithelial cell binding of group A streptococci by lipoteichoic acid on fimbriae denuded of M protein. *J Exp Med*, 143, 759-71.
- BECKMANN, C., WAGGONER, J. D., HARRIS, T. O., TAMURA, G. S. & RUBENS, C. E. (2002) Identification of novel adhesins from Group B streptococci by use of phage display reveals that C5a peptidase mediates fibronectin binding. *Infect Immun*, 70, 2869-76.
- BERGE, A. & BJORCK, L. (1995) Streptococcal cysteine proteinase releases biologically active fragments of streptococcal surface proteins. *J Biol Chem*, 270, 9862-7.
- BERGE, A. & SJOBRING, U. (1993) PAM, a novel plasminogen-binding protein from *Streptococcus pyogenes*. *J Biol Chem*, 268, 25417-24.
- BERGGREN, K., CHERNOKALSKAYA, E., STEINBERG, T. H., KEMPER, C., LOPEZ, M. F., DIWU, Z., HAUGLAND, R. P. & PATTON, W. F. (2000) Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. *Electrophoresis*, 21, 2509-21.
- BETLEY, M. J., BORST, D. W. & REGASSA, L. B. (1992) Staphylococcal enterotoxins, toxic shock syndrome toxin and streptococcal pyrogenic exotoxins: a comparative study of their molecular biology. *Chem Immunol*, 55, 1-35.
- BETSCHER, S. D., BORGIA, S. M., BARG, N. L., LOW, D. E. & DE AZAVEDO, J. C. (1998) Reduced virulence of group A streptococcal Tn916 mutants that do not produce streptolysin S. *Infect Immun*, 66, 1671-9.
- BISNO, A. L. (1979) Alternate complement pathway activation by group A streptococci: role of M-protein. *Infect Immun*, 26, 1172-6.
- BISNO, A. L. (2001) Acute pharyngitis. *N Engl J Med*, 344, 205-11.

- BISNO, A. L., BRITO, M. O. & COLLINS, C. M. (2003) Molecular basis of group A streptococcal virulence. *Lancet Infect Dis*, 3, 191-200.
- BISNO, A. L. & STEVENS, D. L. (1996) Streptococcal infections of skin and soft tissues. *N Engl J Med*, 334, 240-5.
- BRADFORD, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 72, 248-54.
- BRICKER, A. L., CAREY, V. J. & WESSELS, M. R. (2005) Role of NADase in virulence in experimental invasive group A streptococcal infection. *Infect Immun*, 73, 6562-6.
- BRICKER, A. L., CYWES, C., ASHBAUGH, C. D. & WESSELS, M. R. (2002) NAD⁺-glycohydrolase acts as an intracellular toxin to enhance the extracellular survival of group A streptococci. *Mol Microbiol*, 44, 257-69.
- BROUDY, T. B., PANCHOLI, V. & FISCHETTI, V. A. (2002) The in vitro interaction of *Streptococcus pyogenes* with human pharyngeal cells induces a phage-encoded extracellular DNase. *Infect Immun*, 70, 2805-11.
- BUCHANAN, J. T., SIMPSON, A. J., AZIZ, R. K., LIU, G. Y., KRISTIAN, S. A., KOTB, M., FERAMISCO, J. & NIZET, V. (2006) DNase expression allows the pathogen group A *Streptococcus* to escape killing in neutrophil extracellular traps. *Curr Biol*, 16, 396-400.
- BURKITT, H. G. (1993) *Wheater's functional histology : a text and colour atlas*, Edinburgh, Churchill Livinstone.
- BUROVA, L. A., NAGORNEV, V. A., PIGAREVSKY, P. V., GLADILINA, M. M., GAVRILOVA, E. A., SELIVERSTOVA, V. G., TOTOLIAN, A. A., THERN, A. & SCHALEN, C. (2005) Myocardial tissue damage in rabbits injected with

group A streptococci, types M1 and M22. Role of bacterial immunoglobulin G-binding surface proteins. *Apmis*, 113, 21-30.

CAPARON, M. (2001) Pathogenic Mechanisms in streptococcal diseases. IN GROISMAN, A. E. (Ed.) *Principles of bacterial pathogenesis*. London, Academic press.

CAPARON, M. G., STEPHENS, D. S., OLSEN, A. & SCOTT, J. R. (1991) Role of M protein in adherence of group A streptococci. *Infect Immun*, 59, 1811-7.

CARAPETIS, J. R., STEER, A. C., MULHOLLAND, E. K. & WEBER, M. (2005) The global burden of group A streptococcal diseases. *Lancet Infect Dis*, 5, 685-94.

CASH, P., ARGO, E., FORD, L., LAWRIE, L. & MCKENZIE, H. (1999) A proteomic analysis of erythromycin resistance in *Streptococcus pneumoniae*. *Electrophoresis*, 20, 2259-68.

CHAGOYEN, M., CARMONA-SAEZ, P., SHATKAY, H., CARAZO, J. M. & PASCUAL-MONTANO, A. (2006) Discovering semantic features in the literature: a foundation for building functional associations. *BMC Bioinformatics*, 7, 41.

CHAMRAD, D. C., KORTING, G., STUHLER, K., MEYER, H. E., KLOSE, J. & BLUGGEL, M. (2004) Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data. *Proteomics*, 4, 619-28.

CHAUSSEE, M. S., AJDIC, D. & FERRETTI, J. J. (1999) The *rgg* gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SPE B production. *Infect Immun*, 67, 1715-22.

CHAUSSEE, M. A., CALLEGARI, E. A. & CHAUSSEE, M. S. (2004) Rgg regulates growth phase-dependent expression of proteins associated with secondary

metabolism and stress in *Streptococcus pyogenes*. *Journal of Bacteriology*, 186, 7091-7099.

CHAUSSEE, M. S., WATSON, R. O., SMOOT, J. C. & MUSSER, J. M. (2001)

Identification of Rgg-regulated exoproteins of *Streptococcus pyogenes*. *Infect Immun*, 69, 822-31.

CHIANG-NI, C., WANG, C. H., TSAI, P. J., CHUANG, W. J., LIN, Y. S., LIN, M. T.,

LIU, C. C. & WU, J. J. (2006) Streptococcal pyrogenic exotoxin B causes mitochondria damage to polymorphonuclear cells preventing phagocytosis of group A streptococcus. *Med Microbiol Immunol (Berl)*, 195, 55-63.

CHOE, L. H. & LEE, K. H. (2003) Quantitative and qualitative measure of

intralaboratory two-dimensional protein gel reproducibility and the effects of sample preparation, sample load, and image analysis. *Electrophoresis*, 24, 3500-7.

CLEARY, P. P. & LARKIN, A. (1979) Hyaluronic acid capsule: strategy for oxygen

resistance in group A streptococci. *J Bacteriol*, 140, 1090-7.

CLEARY, P. P., PRAHBU, U., DALE, J. B., WEXLER, D. E. & HANDLEY, J. (1992)

Streptococcal C5a peptidase is a highly specific endopeptidase. *Infect Immun*, 60, 5219-23.

CODD, A. A., RICHARDSON, I. R. & ANDREW, N. (1998) Lenticules for the

control of quantitative methods in food microbiology. *J Appl Microbiol*, 85, 913-7.

COLE, J. N., RAMIREZ, R. D., CURRIE, B. J., CORDWELL, S. J., DJORDJEVIC, S.

P. & WALKER, M. J. (2005) Surface analyses and immune reactivities of major cell wall-associated proteins of group a streptococcus. *Infect Immun*, 73, 3137-46.

- COLLIN, M. & OLSEN, A. (2003) Extracellular enzymes with immunomodulating activities: variations on a theme in *Streptococcus pyogenes*. *Infect Immun*, 71, 2983-92.
- COURTNEY, H., OFEK, I., SIMPSON, W. A. & BEACHEY, E. H. (1981) Characterization of lipoteichoic acid binding to polymorphonuclear leukocytes of human blood. *Infect Immun*, 32, 625-31.
- COURTNEY, H. S., BRONZE, M. S., DALE, J. B. & HASTY, D. L. (1994a) Analysis of the role of M24 protein in group A streptococcal adhesion and colonization by use of omega-interposon mutagenesis. *Infect Immun*, 62, 4868-73.
- COURTNEY, H. S. & HASTY, D. L. (1991) Aggregation of group A streptococci by human saliva and effect of saliva on streptococcal adherence to host cells. *Infect Immun*, 59, 1661-6.
- COURTNEY, H. S., HASTY, D. L., LI, Y., CHIANG, H. C., THACKER, J. L. & DALE, J. B. (1999) Serum opacity factor is a major fibronectin-binding protein and a virulence determinant of M type 2 *Streptococcus pyogenes*. *Mol Microbiol*, 32, 89-98.
- COURTNEY, H. S., LI, Y., DALE, J. B. & HASTY, D. L. (1994b) Cloning, sequencing, and expression of a fibronectin/fibrinogen-binding protein from group A streptococci. *Infect Immun*, 62, 3937-46.
- COURTNEY, H. S., VON HUNOLSTEIN, C., DALE, J. B., BRONZE, M. S., BEACHEY, E. H. & HASTY, D. L. (1992) Lipoteichoic acid and M protein: dual adhesins of group A streptococci. *Microb Pathog*, 12, 199-208.
- CRATER, D. L. & VAN DE RIJN, I. (1995) Hyaluronic acid synthesis operon (has) expression in group A streptococci. *J Biol Chem*, 270, 18452-8.
- CU, G. A., MEZZANO, S., BANNAN, J. D. & ZABRISKIE, J. B. (1998) Immunohistochemical and serological evidence for the role of streptococcal

proteinase in acute post-streptococcal glomerulonephritis. *Kidney Int*, 54, 819-26.

CUNNINGHAM, C. M. & WATSON, D. W. (1978) Suppression of antibody response by group A streptococcal pyrogenic exotoxin and characterization of the cells involved. *Infect Immun*, 19, 470-6.

CUNNINGHAM, M. W. (2000) Pathogenesis of group A streptococcal infections. *Clin Microbiol Rev*, 13, 470-511.

CYWES, C., STAMENKOVIC, I. & WESSELS, M. R. (2000) CD44 as a receptor for colonization of the pharynx by group A Streptococcus. *J Clin Invest*, 106, 995-1002.

DALE, J. B., WASHBURN, R. G., MARQUES, M. B. & WESSELS, M. R. (1996) Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. *Infect Immun*, 64, 1495-501.

DALTON, T. L. & SCOTT, J. R. (2004) CovS inactivates CovR and is required for growth under conditions of general stress in Streptococcus pyogenes. *J Bacteriol*, 186, 3928-37.

DARMSTADT, G. L., MENTELE, L., PODBIELSKI, A. & RUBENS, C. E. (2000) Role of group A streptococcal virulence factors in adherence to keratinocytes. *Infect Immun*, 68, 1215-21.

DIXON, S., HASWELL, M., HARRINGTON, D. & SUTCLIFFE, I. C. (2001) Surface immunolocalisation of HPr in the equine pathogen Streptococcus equi. *Syst Appl Microbiol*, 24, 486-9.

DRYLA, A., PRUSTOMERSKY, S., GELBMANN, D., HANNER, M., BETTINGER, E., KOCSIS, B., KUSTOS, T., HENICS, T., MEINKE, A. & NAGY, E. (2005) Comparison of antibody repertoires against Staphylococcus aureus in healthy

- individuals and in acutely infected patients. *Clin Diagn Lab Immunol*, 12, 387-98.
- DUTT, M. J. & LEE, K. H. (2000) Proteomic analysis. *Curr Opin Biotechnol*, 11, 176-9.
- ELLEN, R. P. & GIBBONS, R. J. (1972) M protein-associated adherence of *Streptococcus pyogenes* to epithelial surfaces: prerequisite for virulence. *Infect Immun*, 5, 826-30.
- ENG, J. K., MCCORMACK, A. L. & YATES, J. R. (1994) An Approach to Correlate Tandem Mass-Spectral Data of Peptides with Amino-Acid-Sequences in a Protein Database. *J Am Soc Mass Spectrom*, 5, 976-989.
- ERIKSSON, A. & NORGREN, M. (2003) Cleavage of antigen-bound immunoglobulin G by SpeB contributes to streptococcal persistence in opsonizing blood. *Infect Immun*, 71, 211-7.
- FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246, 64-71.
- FERRETTI, J. J., MCSHAN, W. M., AJDIC, D., SAVIC, D. J., SAVIC, G., LYON, K., PRIMEAUX, C., SEZATE, S., SUVOROV, A. N., KENTON, S., LAI, H. S., LIN, S. P., QIAN, Y., JIA, H. G., NAJAR, F. Z., REN, Q., ZHU, H., SONG, L., WHITE, J., YUAN, X., CLIFTON, S. W., ROE, B. A. & MCLAUGHLIN, R. (2001) Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A*, 98, 4658-63.
- FISCHETTI, V. A. (1989) Streptococcal M protein: molecular design and biological behavior. *Clin Microbiol Rev*, 2, 285-314.
- FLEISCHMANN, R. D., ADAMS, M. D., WHITE, O., CLAYTON, R. A., KIRKNESS, E. F., KERLAVAGE, A. R., BULT, C. J., TOMB, J. F.,

- DOUGHERTY, B. A., MERRICK, J. M. & ET AL. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 269, 496-512.
- FOLEY, M. J. & WOOD, W. B., JR. (1959) Studies on the pathogenicity of group A streptococci. II. The antiphagocytic effects of the M protein and the capsular gel. *J Exp Med*, 110, 617-28.
- FONTAINE, M. C., LEE, J. J. & KEHOE, M. A. (2003) Combined contributions of streptolysin O and streptolysin S to virulence of serotype M5 *Streptococcus pyogenes* strain manfredo. *Infection and Immunity*, 71, 3857-3865.
- FRITZ, E., HEDEN, L. O. & LINDAHL, G. (1989) Extensive sequence homology between IgA receptor and M proteins in *Streptococcus pyogenes*. *Mol Microbiol*, 3, 1111-9.
- GAHAN, C. G. & HILL, C. (2000) The use of listeriolysin to identify in vivo induced genes in the gram-positive intracellular pathogen *Listeria monocytogenes*. *Mol Microbiol*, 36, 498-507.
- GALINIER, A., HAIECH, J., KILHOFFER, M. C., JAQUINOD, M., STULKE, J., DEUTSCHER, J. & MARTIN-VERSTRAETE, I. (1997) The *Bacillus subtilis* *crh* gene encodes a HPr-like protein involved in carbon catabolite repression. *Proc Natl Acad Sci U S A*, 94, 8439-44.
- GERLACH, D., KNOLL, H., KOHLER, W., OZEGOWSKI, J. H. & HRIBALOVA, V. (1983) Isolation and characterization of erythrogenic toxins. V. Communication: identity of erythrogenic toxin type B and streptococcal proteinase precursor. *Zentralbl Bakteriol Mikrobiol Hyg [A]*, 255, 221-33.
- GINSBURG, I. (1999) Is streptolysin S of group A streptococci a virulence factor? *Apmis*, 107, 1051-9.

- GLISH, G. & VACHET, R. (2003) The basics of mass spectrometry in the twenty-first century. *Nature*, 2, 140-150.
- GOMI, H., HOZUMI, T., HATTORI, S., TAGAWA, C., KISHIMOTO, F. & BJORCK, L. (1990) The gene sequence and some properties of protein H. A novel IgG-binding protein. *J Immunol*, 144, 4046-52.
- GORG, A., OBERMAIER, C., BOGUTH, G., HARDER, A., SCHEIBE, B., WILDGRUBER, R. & WEISS, W. (2000) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*, 21, 1037-53.
- GORG, A., POSTEL, W. & GUNTHER, S. (1988) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*, 9, 531-46.
- GOSHORN, S. C. & SCHLIEVERT, P. M. (1989) Bacteriophage association of streptococcal pyrogenic exotoxin type C. *J Bacteriol*, 171, 3068-73.
- GOVORUN, V. M. & ARCHAKOV, A. I. (2002) Proteomic technologies in modern biomedical science. *Biochemistry (Mosc)*, 67, 1109-23.
- GRAHAM, M. R., SMOOT, L. M., LEI, B. & MUSSER, J. M. (2001) Toward a genome-scale understanding of group A Streptococcus pathogenesis. *Curr Opin Microbiol*, 4, 65-70.
- GRAHAM, M. R., VIRTANEVA, K., PORCELLA, S. F., BARRY, W. T., GOWEN, B. B., JOHNSON, C. R., WRIGHT, F. A. & MUSSER, J. M. (2005) Group A Streptococcus transcriptome dynamics during growth in human blood reveals bacterial adaptive and survival strategies. *Am J Pathol*, 166, 455-65.
- GREEN, N. M., ZHANG, S., PORCELLA, S. F., NAGIEC, M. J., BARBIAN, K. D., BERES, S. B., LEFEBVRE, R. B. & MUSSER, J. M. (2005) Genome sequence of a serotype M28 strain of group a streptococcus: potential new

insights into puerperal sepsis and bacterial disease specificity. *J Infect Dis*, 192, 760-70.

GREENWOOD, D., SLACK, R. C. B. & PEUTHERER, J. F. (1992) *Medical microbiology : a guide to microbial infections : pathogenesis, immunity, laboratory diagnosis and control*, Churchill Livingstone, Edinburgh.

GUBBA, S., LOW, D. E. & MUSSER, J. M. (1998) Expression and characterization of group A Streptococcus extracellular cysteine protease recombinant mutant proteins and documentation of seroconversion during human invasive disease episodes. *Infect Immun*, 66, 765-70.

GUYRE, P. M., CAMPBELL, A. S., KNIFFIN, W. D. & FANGER, M. W. (1990) Monocytes and polymorphonuclear neutrophils of patients with streptococcal pharyngitis express increased numbers of type I IgG Fc receptors. *J Clin Invest*, 86, 1892-6.

GYGI, S. P., ROCHON, Y., FRANZA, B. R. & AEBERSOLD, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol*, 19, 1720-30.

HAKANSSON, A., BENTLEY, C. C., SHAKHNOVIC, E. A. & WESSELS, M. R. (2005) Cytolysin-dependent evasion of lysosomal killing. *Proc Natl Acad Sci USA*, 102, 5192-7.

HALPERIN, S. A., FERRIERI, P., GRAY, E. D., KAPLAN, E. L. & WANNAMAKER, L. W. (1987) Antibody response to bacteriophage hyaluronidase in acute glomerulonephritis after group A streptococcal infection. *J Infect Dis*, 155, 253-61.

HAMER, L., DEZWAAN, T. M., MONTENEGRO-CHAMORRO, M. V., FRANK, S. A. & HAMER, J. E. (2001) Recent advances in large-scale transposon mutagenesis. *Curr Opin Chem Biol*, 5, 67-73.

- HANSKI, E. & CAPARON, M. (1992) Protein F, a fibronectin-binding protein, is an adhesin of the group A streptococcus *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A*, 89, 6172-6.
- HAVA, D. L. & CAMILLI, A. (2002) Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol*, 45, 1389-406.
- HOLM, S. E., NORRBY, A., BERGHOLM, A. M. & NORGREN, M. (1992) Aspects of pathogenesis of serious group A streptococcal infections in Sweden, 1988-1989. *J Infect Dis*, 166, 31-7.
- HOMER, K. A., GROOTVELD, M. C., HAWKES, J., NAUGHTON, D. P. & BEIGHTON, D. (1994) Degradation of hyaluronate by *Streptococcus intermedius* strain UNS 35. *J Med Microbiol*, 41, 414-22.
- HOMER, K. A., PATEL, R. & BEIGHTON, D. (1993) Effects of N-acetylglucosamine on carbohydrate fermentation by *Streptococcus mutans* NCTC 10449 and *Streptococcus sobrinus* SL-1. *Infect Immun*, 61, 295-302.
- HOSEIN, B., MCCARTY, M. & FISCHETTI, V. A. (1979) Amino acid sequence and physicochemical similarities between streptococcal M protein and mammalian tropomyosin. *Proc Natl Acad Sci U S A*, 76, 3765-8.
- HOSTETTER, M. K. (1986) Serotypic variations among virulent pneumococci in deposition and degradation of covalently bound C3b: implications for phagocytosis and antibody production. *J Infect Dis*, 153, 682-93.
- HOUTMAN, R., KRIJGSVELD, J., KOOL, M., ROMIJN, E. P., REDEGELD, F. A., NIJKAMP, F. P., HECK, A. J. & HUMPHERY-SMITH, I. (2003) Lung proteome alterations in a mouse model for nonallergic asthma. *Proteomics*, 3, 2008-18.
- HUANG, T. T., MALKE, H. & FERRETTI, J. J. (1989) Heterogeneity of the streptokinase gene in group A streptococci. *Infect Immun*, 57, 502-6.

- HUBER, M., TRATTNIG, S. & LINTNER, F. (2000) Anatomy, biochemistry, and physiology of articular cartilage. *Invest Radiol*, 35, 573-80.
- HUGHES, T. R., MARTON, M. J., JONES, A. R., ROBERTS, C. J., STOUGHTON, R., ARMOUR, C. D., BENNETT, H. A., COFFEY, E., DAI, H., HE, Y. D., KIDD, M. J., KING, A. M., MEYER, M. R., SLADE, D., LUM, P. Y., STEPANIANTS, S. B., SHOEMAKER, D. D., GACHOTTE, D., CHAKRABURTTY, K., SIMON, J., BARD, M. & FRIEND, S. H. (2000) Functional discovery via a compendium of expression profiles. *Cell*, 102, 109-26.
- HUSMANN, L. K., YUNG, D. L., HOLLINGSHEAD, S. K. & SCOTT, J. R. (1997) Role of putative virulence factors of *Streptococcus pyogenes* in mouse models of long-term throat colonization and pneumonia. *Infect Immun*, 65, 1422-30.
- HYNES, W. (2004) Virulence factors of the group a streptococci and genes that regulate their expression. *Front Biosci*, 9, 3399-433.
- HYNES, W. L. & FERRETTI, J. J. (1989) Sequence analysis and expression in *Escherichia coli* of the hyaluronidase gene of *Streptococcus pyogenes* bacteriophage H4489A. *Infect Immun*, 57, 533-9.
- HYNES, W. L., HANCOCK, L. & FERRETTI, J. J. (1995) Analysis of a second bacteriophage hyaluronidase gene from *Streptococcus pyogenes*: evidence for a third hyaluronidase involved in extracellular enzymatic activity. *Infect Immun*, 63, 3015-20.
- HYNES, W. L. & WALTON, S. L. (2000) Hyaluronidases of Gram-positive bacteria. *FEMS Microbiol Lett*, 183, 201-7.
- HYTONEN, J., HAATAJA, S., GERLACH, D., PODBIELSKI, A. & FINNE, J. (2001) The SpeB virulence factor of *Streptococcus pyogenes*, a multifunctional

secreted and cell surface molecule with strepadhesin, laminin-binding and cysteine protease activity. *Mol Microbiol*, 39, 512-9.

- IIDA, K., SEKI, M., SAITO, M., KAWAMURA, Y., KAJIWARA, H. & YOSHIDA, S. (2006) Capsule of *Streptococcus pyogenes* is essential for delayed death of mice in a model of streptococcal toxic shock syndrome. *Microbiol Immunol*, 50, 127-30.
- IKEBE, T., ENDOH, M. & WATANABE, H. (2005) Increased expression of the *ska* gene in emm49-genotyped *Streptococcus pyogenes* strains isolated from patients with severe invasive streptococcal infections. *Jpn J Infect Dis*, 58, 272-5.
- IYER, R., BALIGA, N. S. & CAMILLI, A. (2005) Catabolite control protein A (CcpA) contributes to virulence and regulation of sugar metabolism in *Streptococcus pneumoniae*. *J Bacteriol*, 187, 8340-9.
- Jl, Y., MCLANDSBOROUGH, L., KONDAGUNTA, A. & CLEARY, P. P. (1996) C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. *Infect Immun*, 64, 503-10.
- JOHANSSON, B. P., LEVANDER, F., VON PAWEL-RAMMINGEN, U., BERGGARD, T., BJORCK, L. & JAMES, P. (2005) The protein expression of *Streptococcus pyogenes* is significantly influenced by human plasma. *Journal of Proteome Research*, 4, 2302-2311.
- JOHANSSON, H. M., MORGELIN, M. & FRICK, I. M. (2004) Protein FOG--a streptococcal inhibitor of neutrophil function. *Microbiology*, 150, 4211-21.
- JOHNSON, D. R. & KAPLAN, E. L. (1993) A review of the correlation of T-agglutination patterns and M-protein typing and opacity factor production in the identification of group A streptococci. *J Med Microbiol*, 38, 311-5.
- JOHNSON, E. K. (1991) *Histology and cell biology*, Baltimore, Williams and Wilkins.

- JOHNSON, L. P. & SCHLIEVERT, P. M. (1984) Group A streptococcal phage T12 carries the structural gene for pyrogenic exotoxin type A. *Mol Gen Genet*, 194, 52-6.
- JONES, A. L., KNOLL, K. M. & RUBENS, C. E. (2000) Identification of *Streptococcus agalactiae* virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. *Mol Microbiol*, 37, 1444-55.
- KANEHISA, M. & BORK, P. (2003) Bioinformatics in the post-sequence era. *Nat Genet*, 33 Suppl, 305-10.
- KANSAL, R. G., MCGEER, A., LOW, D. E., NORRBY-TEGLUND, A. & KOTB, M. (2000) Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. *Infect Immun*, 68, 6362-9.
- KAPUR, V., MAJESKY, M. W., LI, L. L., BLACK, R. A. & MUSSER, J. M. (1993a) Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A*, 90, 7676-80.
- KAPUR, V., TOPOUZIS, S., MAJESKY, M. W., LI, L. L., HAMRICK, M. R., HAMILL, R. J., PATTI, J. M. & MUSSER, J. M. (1993b) A conserved *Streptococcus pyogenes* extracellular cysteine protease cleaves human fibronectin and degrades vitronectin. *Microb Pathog*, 15, 327-46.
- KAWABATA, S., KUWATA, H., NAKAGAWA, I., MORIMATSU, S., SANO, K. & HAMADA, S. (1999) Capsular hyaluronic acid of group A streptococci hampers their invasion into human pharyngeal epithelial cells. *Microb Pathog*, 27, 71-80.

- KELLAM, P. (2001) Post-genomic virology: the impact of bioinformatics, microarrays and proteomics on investigating host and pathogen interactions. *Reviews in Medical Virology*, 11, 313-329.
- KOLKER, E., HIGDON, R. & HOGAN, J. M. (2006) Protein identification and expression analysis using mass spectrometry. *Trends in Microbiology*, 14, 229-235.
- KOTARSKY, H., GUSTAFSSON, M., SVENSSON, H. G., ZIPFEL, P. F., TRUEDSSON, L. & SJOBRING, U. (2001) Group A streptococcal phagocytosis resistance is independent of complement factor H and factor H-like protein 1 binding. *Mol Microbiol*, 41, 817-26.
- KRAUSE, J. W. (1996) *Essentials of human histology*, London, Little, brown and company.
- KUETTNER, K. E. (1992) Biochemistry of articular cartilage in health and disease. *Clin Biochem*, 25, 155-63.
- LA PENTA, D., ZHANG, X. P. & CLEARY, P. P. (1994) Streptococcus pyogenes type IIa IgG Fc receptor expression is co-ordinately regulated with M protein and streptococcal C5a peptidase. *Mol Microbiol*, 12, 873-9.
- LANCEFIELD, R. C. (1962) Current knowledge of type-specific M antigens of group A streptococci. *J Immunol*, 89, 307-13.
- LAU, G. W., HAATAJA, S., LONETTO, M., KENSIT, S. E., MARRA, A., BRYANT, A. P., MCDEVITT, D., MORRISON, D. A. & HOLDEN, D. W. (2001) A functional genomic analysis of type 3 Streptococcus pneumoniae virulence. *Mol Microbiol*, 40, 555-71.
- LAURENT, T. C. & FRASER, J. R. (1992) Hyaluronan. *Faseb J*, 6, 2397-404.

- LEI, B., MACKIE, S., LUKOMSKI, S. & MUSSER, J. M. (2000) Identification and immunogenicity of group A Streptococcus culture supernatant proteins. *Infect Immun*, 68, 6807-18.
- LEONARD, B. A., WOISCHNIK, M. & PODBIELSKI, A. (1998) Production of stabilized virulence factor-negative variants by group A streptococci during stationary phase. *Infect Immun*, 66, 3841-7.
- LI, S., KELLY, S. J., LAMANI, E., FERRARONI, M. & JEDRZEJAS, M. J. (2000) Structural basis of hyaluronan degradation by Streptococcus pneumoniae hyaluronate lyase. *Embo J*, 19, 1228-40.
- LI, Z., PLOPLIS, V. A., FRENCH, E. L. & BOYLE, M. D. (1999) Interaction between group A streptococci and the plasmin(ogen) system promotes virulence in a mouse skin infection model. *J Infect Dis*, 179, 907-14.
- LICALSI, C., CROCENZI, T. S., FREIRE, E. & ROSEMAN, S. (1991) Sugar transport by the bacterial phosphotransferase system. Structural and thermodynamic domains of enzyme I of Salmonella typhimurium. *J Biol Chem*, 266, 19519-27.
- LIMBAGO, B., PENUMALLI, V., WEINRICK, B. & SCOTT, J. R. (2000) Role of streptolysin O in a mouse model of invasive group A streptococcal disease. *Infect Immun*, 68, 6384-90.
- LINDAHL, U. & HOOK, M. (1978) Glycosaminoglycans and their binding to biological macromolecules. *Annu Rev Biochem*, 47, 385-417.
- LIOLIOS, K., TAVERNARAKIS, N., HUGENHOLTZ, P. & KYRPIDES, N. C. (2006) The Genomes On Line Database (GOLD) v.2: a monitor of genome projects worldwide. *Nucleic Acids Res*, 34, D332-4.
- LONG, X. H., ZHU, J. W., MO, Z. H., FENG, S., CHENG, G., ZHOU, X. W., ZHANG, Y. Z. & YANG, P. Y. (2006) Development of an effective sample

preparation approach for proteomic analysis of silkworm eggs using two-dimensional gel electrophoresis and mass spectrometry. *J Chromatogr A*.

LOTTENBERG, R., MINNING-WENZ, D. & BOYLE, M. D. (1994) Capturing host plasmin(ogen): a common mechanism for invasive pathogens? *Trends Microbiol*, 2, 20-4.

LUKOMSKI, S., NAKASHIMA, K., ABDI, I., CIPRIANO, V. J., IRELAND, R. M., REID, S. D., ADAMS, G. G. & MUSSER, J. M. (2000) Identification and characterization of the scl gene encoding a group A Streptococcus extracellular protein virulence factor with similarity to human collagen. *Infect Immun*, 68, 6542-53.

LYON, W. R., GIBSON, C. M. & CAPARON, M. G. (1998) A role for trigger factor and an rgg-like regulator in the transcription, secretion and processing of the cysteine proteinase of Streptococcus pyogenes. *Embo J*, 17, 6263-75.

MADDEN, J. C., RUIZ, N. & CAPARON, M. (2001) Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria. *Cell*, 104, 143-52.

MAHAN, J. D., HEBERT, L. A., MCALLISTER, C., BIRMINGHAM, D. J., SHEN, X. P., COSIO, F. G. & BRANDT, J. (1993) Platelet involvement in experimental immune complex-mediated glomerulonephritis in the nonhuman primate. *Kidney Int*, 44, 716-25.

MAHAN, M. J., TOBIAS, J. W., SLAUCH, J. M., HANNA, P. C., COLLIER, R. J. & MEKALANOS, J. J. (1995) Antibiotic-based selection for bacterial genes that are specifically induced during infection of a host. *Proc Natl Acad Sci U S A*, 92, 669-73.

MALKE, H., STEINER, K., MCSHAN, W. M. & FERRETTI, J. J. (2006) Linking the nutritional status of Streptococcus pyogenes to alteration of transcriptional gene expression: The action of CodY and RelA. *Int J Med Microbiol*.

- MARRA, A., ASUNDI, J., BARTILSON, M., LAWSON, S., FANG, F., CHRISTINE, J., WIESNER, C., BRIGHAM, D., SCHNEIDER, W. P. & HROMOCKYJ, A. E. (2002) Differential fluorescence induction analysis of *Streptococcus pneumoniae* identifies genes involved in pathogenesis. *Infect Immun*, 70, 1422-33.
- MATSUKA, Y. V., PILLAI, S., GUBBA, S., MUSSER, J. M. & OLMSTED, S. B. (1999) Fibrinogen cleavage by the *Streptococcus pyogenes* extracellular cysteine protease and generation of antibodies that inhibit enzyme proteolytic activity. *Infect Immun*, 67, 4326-33.
- MERRELL, D. S. & CAMILLI, A. (2000) Detection and analysis of gene expression during infection by in vivo expression technology. *Philos Trans R Soc Lond B Biol Sci*, 355, 587-99.
- MILLER, G. L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry*, 426-428.
- MIMS, C., DOCKRELL, H., GOERING, R., ROITT, I., WAKELIN, D. & ZUCKERMAN, M. (1998) *Medical Microbiology*, Mosby, USA.
- MISCHE, S. M., MANJULA, B. N. & FISCHETTI, V. A. (1987) Relation of streptococcal M protein with human and rabbit tropomyosin: the complete amino acid sequence of human cardiac alpha tropomyosin, a highly conserved contractile protein. *Biochem Biophys Res Commun*, 142, 813-8.
- MIYOSHI-AKIYAMA, T., TAKAMATSU, D., KOYANAGI, M., ZHAO, J., IMANISHI, K. & UCHIYAMA, T. (2005) Cytocidal effect of *Streptococcus pyogenes* on mouse neutrophils in vivo and the critical role of streptolysin S. *J Infect Dis*, 192, 107-16.
- MOSES, A. E., WESSELS, M. R., ZALCMAN, K., ALBERTI, S., NATANSON-YARON, S., MENES, T. & HANSKI, E. (1997) Relative contributions of

hyaluronic acid capsule and M protein to virulence in a mucoid strain of the group A Streptococcus. *Infect Immun*, 65, 64-71.

MUSSER, J. M. & DELEO, F. R. (2005) Toward a genome-wide systems biology analysis of host-pathogen interactions in group A Streptococcus. *American Journal of Pathology*, 167, 1461-1472.

NA, J. G., PINTO, I. & HAMPSEY, M. (1992) Isolation and characterization of SUA5, a novel gene required for normal growth in *Saccharomyces cerevisiae*. *Genetics*, 131, 791-801.

NAKAMURA, T., HASEGAWA, T., TORII, K., HASEGAWA, Y., SHIMOKATA, K. & OHTA, M. (2004) Two-dimensional gel electrophoresis analysis of the abundance of virulent exoproteins of group A streptococcus caused by environmental changes. *Arch Microbiol*, 181, 74-81.

NAKATA, M., PODBIELSKI, A. & KREIKEMEYER, B. (2005) MsmR, a specific positive regulator of the *Streptococcus pyogenes* FCT pathogenicity region and cytolysin-mediated translocation system genes. *Mol Microbiol*, 57, 786-803.

NANNINI, E. C., TENG, F., SINGH, K. V. & MURRAY, B. E. (2005) Decreased virulence of a gls24 mutant of *Enterococcus faecalis* OG1RF in an experimental endocarditis model. *Infect Immun*, 73, 7772-4.

NATANSON, S., SELA, S., MOSES, A. E., MUSSER, J. M., CAPARON, M. G. & HANSKI, E. (1995) Distribution of fibronectin-binding proteins among group A streptococci of different M types. *J Infect Dis*, 171, 871-8.

NEELEMAN, C., GEELEN, S. P., AERTS, P. C., DAHA, M. R., MOLLNES, T. E., ROORD, J. J., POSTHUMA, G., VAN DIJK, H. & FLEER, A. (1999) Resistance to both complement activation and phagocytosis in type 3 pneumococci is mediated by the binding of complement regulatory protein factor H. *Infect Immun*, 67, 4517-24.

- NYBERG, P., RASMUSSEN, M., VON PAWEL-RAMMINGEN, U. & BJORCK, L. (2004) SpeB modulates fibronectin-dependent internalization of *Streptococcus pyogenes* by efficient proteolysis of cell-wall-anchored protein F1. *Microbiology*, 150, 1559-69.
- OEHMEN, C. S., STRAATSMA, T. P., ANDERSON, G. A., ORR, G., WEBB-ROBERTSON, B. J. M., TAYLOR, R. C., MOONEY, R. W., BAXTER, D. J., JONES, D. R. & DIXON, D. A. (2006) New challenges facing integrative biological science in the post-genomic era. *Journal of Biological Systems*, 14, 275-293.
- O'FARRELL, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, 250, 4007-21.
- OHKUNI, H., TODOME, Y., WATANABE, Y., ISHIKAW, T., TAKAHASHI, H., KANNARI, Y., KATO, H., UCHIYAMA, T., SAITO, H., FISCHETTI, V. A. & ZABRISKIE, J. B. (2004) Studies of recombinant streptococcal pyrogenic exotoxin B/cysteine protease (rSPE B/SCP) in the skin of guinea pigs & the release of histamine from cultured mast cells & basophilic leukocytes. *Indian J Med Res*, 119 Suppl, 33-6.
- OKADA, N., LISZEWSKI, M. K., ATKINSON, J. P. & CAPARON, M. (1995) Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of the group A streptococcus. *Proc Natl Acad Sci U S A*, 92, 2489-93.
- OKADA, N., PENTLAND, A. P., FALK, P. & CAPARON, M. G. (1994) M protein and protein F act as important determinants of cell-specific tropism of *Streptococcus pyogenes* in skin tissue. *J Clin Invest*, 94, 965-77.
- OKAMOTO, S., KAWABATA, S., TERAOKA, Y., FUJITAKA, H., OKUNO, Y. & HAMADA, S. (2004) The *Streptococcus pyogenes* capsule is required for adhesion of bacteria to virus-infected alveolar epithelial cells and lethal bacterial-viral superinfection. *Infect Immun*, 72, 6068-75.

- OLIVIER, C. (2000) Rheumatic fever--is it still a problem? *J Antimicrob Chemother*, 45 Suppl, 13-21.
- PAIK, S., SENTY, L., DAS, S., NOE, J. C., MUNRO, C. L. & KITTEN, T. (2005) Identification of virulence determinants for endocarditis in *Streptococcus sanguinis* by signature-tagged mutagenesis. *Infect Immun*, 73, 6064-74.
- PANCHOLI, V. & FISCHETTI, V. A. (1992) A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate-dehydrogenase with multiple binding activity. *J Exp Med*, 176, 415-26.
- PANCHOLI, V. & FISCHETTI, V. A. (1993) Glyceraldehyde-3-phosphate dehydrogenase on the surface of group A streptococci is also an ADP-ribosylating enzyme. *Proc Natl Acad Sci U S A*, 90, 8154-8.
- PANCHOLI, V. & FISCHETTI, V. A. (1997) Cell-to-cell signalling between group A streptococci and pharyngeal cells. Role of streptococcal surface dehydrogenase (SDH). *Adv Exp Med Biol*, 418, 499-504.
- PENNINGTON, S. & DUNN, M. (2001) *Proteomics: from protein sequence to function*, Oxford, BIOS Scientific Publishers Ltd.
- PEREZ-CABALLERO, D., ALBERTI, S., VIVANCO, F., SANCHEZ-CORRAL, P. & RODRIGUEZ DE CORDOBA, S. (2000) Assessment of the interaction of human complement regulatory proteins with group A *Streptococcus*. Identification of a high-affinity group A *Streptococcus* binding site in FHL-1. *Eur J Immunol*, 30, 1243-53.
- PEREZ-CASAL, J., OKADA, N., CAPARON, M. G. & SCOTT, J. R. (1995) Role of the conserved C-repeat region of the M protein of *Streptococcus pyogenes*. *Mol Microbiol*, 15, 907-16.

- PETERSON, P. K., SCHMELING, D., CLEARY, P. P., WILKINSON, B. J., KIM, Y. & QUIE, P. G. (1979) Inhibition of alternative complement pathway opsonization by group A streptococcal M protein. *J Infect Dis*, 139, 575-85.
- PODBIELSKI, A., SCHNITZLER, N., BEYHS, P. & BOYLE, M. D. (1996) M-related protein (Mrp) contributes to group A streptococcal resistance to phagocytosis by human granulocytes. *Mol Microbiol*, 19, 429-41.
- POLISSI, A., PONTIGGIA, A., FEGER, G., ALTIERI, M., MOTTI, H., FERRARI, L. & SIMON, D. (1998) Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infect Immun*, 66, 5620-9.
- PURUSHOTHAMAN, S. S., PARK, H. S. & CLEARY, P. P. (2004) Promotion of fibronectin independent invasion by C5a peptidase into epithelial cells in group A *Streptococcus*. *Indian J Med Res*, 119 Suppl, 44-7.
- RABILLOUD, T., VALETTE, C. & LAWRENCE, J. J. (1994) Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension. *Electrophoresis*, 15, 1552-8.
- RAEDER, R., HAROKOPAKIS, E., HOLLINGSHEAD, S. & BOYLE, M. D. (2000) Absence of SpeB production in virulent large capsular forms of group A streptococcal strain 64. *Infect Immun*, 68, 744-51.
- RAEDER, R., WOISCHNIK, M., PODBIELSKI, A. & BOYLE, M. D. (1998) A secreted streptococcal cysteine protease can cleave a surface-expressed M1 protein and alter the immunoglobulin binding properties. *Res Microbiol*, 149, 539-48.
- RAMAGLI, L. S. & RODRIGUEZ, L. V. (1985) Quantitation of Microgram Amounts of Protein in Two-Dimensional Polyacrylamide-Gel Electrophoresis Sample Buffer. *Electrophoresis*, 6, 559-563.

- RATNAYAKE-LECAMWASAM, M., SERROR, P., WONG, K. W. & SONENSHEIN, A. L. (2001) *Bacillus subtilis* CodY represses early-stationary-phase genes by sensing GTP levels. *Genes Dev*, 15, 1093-103.
- REDIERS, H., RAINEY, P. B., VANDERLEYDEN, J. & DE MOT, R. (2005) Unraveling the secret lives of bacteria: use of in vivo expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression. *Microbiol Mol Biol Rev*, 69, 217-61.
- REED, G. L., LIN, L. F., PARHAMI-SEREN, B. & KUSSIE, P. (1995) Identification of a plasminogen binding region in streptokinase that is necessary for the creation of a functional streptokinase-plasminogen activator complex. *Biochemistry*, 34, 10266-71.
- RENZONE, G., D'AMBROSIO, C., ARENA, S., RULLO, R., LEDDA, L., FERRARA, L. & SCALONI, A. (2005) Differential proteomic analysis in the study of prokaryotes stress resistance. *Ann Ist Super Sanita*, 41, 459-68.
- REZCALLAH, M. S., BOYLE, M. D. & SLEDJESKI, D. D. (2004) Mouse skin passage of *Streptococcus pyogenes* results in increased streptokinase expression and activity. *Microbiology*, 150, 365-71.
- RINGDAHL, U., SVENSSON, M., WISTEDT, A. C., RENNE, T., KELLNER, R., MULLER-ESTERL, W. & SJOBRING, U. (1998) Molecular co-operation between protein PAM and streptokinase for plasmin acquisition by *Streptococcus pyogenes*. *J Biol Chem*, 273, 6424-30.
- ROBERSON, E. B. & FIRESTONE, M. K. (1992) Relationship between Desiccation and Exopolysaccharide Production in a Soil *Pseudomonas* sp. *Appl Environ Microbiol*, 58, 1284-1291.
- ROBERTS, I. S. (1996) The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu Rev Microbiol*, 50, 285-315.

- ROTTA, J., KRAUSE, R. M., LANCEFIELD, R. C., EVERLY, W. & LACKLAND, H. (1971) New approaches for the laboratory recognition of M types of group A streptococci. *J Exp Med*, 134, 1298-315.
- SALYERS, A. A. & WHITT, D. D. (2002) *Bacterial pathogenesis: A molecular approach*, Washington, USA, ASM Press.
- SCHMIDT, K. H., GUNTHER, E. & COURTNEY, H. S. (1996) Expression of both M protein and hyaluronic acid capsule by group A streptococcal strains results in a high virulence for chicken embryos. *Med Microbiol Immunol (Berl)*, 184, 169-173.
- SCHRAGER, H. M., ALBERTI, S., CYWES, C., DOUGHERTY, G. J. & WESSELS, M. R. (1998) Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of group A Streptococcus to CD44 on human keratinocytes. *J Clin Invest*, 101, 1708-16.
- SCHWARTZ, I. (2000) Microbial genomics: from sequence to function. *Emerg Infect Dis*, 6, 493-5.
- SCHWARZ-LINEK, U., HOOK, M. & POTTS, J. R. (2006) Fibronectin-binding proteins of Gram-positive cocci. *Microbes Infect*.
- SENADHEERA, M. D., GUGGENHEIM, B., SPATAFORA, G. A., HUANG, Y. C., CHOI, J., HUNG, D. C., TREGLOWN, J. S., GOODMAN, S. D., ELLEN, R. P. & CVITKOVITCH, D. G. (2005) A VicRK signal transduction system in Streptococcus mutans affects gtfBCD, gbpB, and ftf expression, biofilm formation, and genetic competence development. *J Bacteriol*, 187, 4064-76.
- SIMPSON, W. A. & BEACHEY, E. H. (1983) Adherence of group A streptococci to fibronectin on oral epithelial cells. *Infect Immun*, 39, 275-9.
- SJOBRING, U., RINGDAHL, U. & RUGGERI, Z. M. (2002) Induction of platelet thrombi by bacteria and antibodies. *Blood*, 100, 4470-7.

- SMITH, H. E., BUIJS, H., DE VRIES, R. R., WISSELINK, H. J., STOCKHOFE-ZURWIEDEN, N. & SMITS, M. A. (2001) Environmentally regulated genes of *Streptococcus suis*: identification by the use of iron-restricted conditions in vitro and by experimental infection of piglets. *Microbiology*, 147, 271-80.
- SMITH, L. N. (2004) Analysis of a family 6 Carbohydrate-Binding Module and three family 69 Hyaluronidases. *School of Applied Sciences*. Newcastle upon Tyne, Northumbria University at Newcastle upon Tyne.
- SMITH, N. L., TAYLOR, E. J., LINDSAY, A. M., CHARNOCK, S. J., TURKENBURG, J. P., DODSON, E. J., DAVIES, G. J. & BLACK, G. W. (2005) Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded beta-helix. *Proc Natl Acad Sci U S A*, 102, 17652-7.
- SMOOT, L. M., SMOOT, J. C., GRAHAM, M. R., SOMERVILLE, G. A., STURDEVANT, D. E., MIGLIACCIO, C. A. L., SYLVA, G. L. & MUSSER, J. M. (2001) Global differential gene expression in response to growth temperature alteration in group A *Streptococcus*. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 10416-10421.
- SRISKANDAN, S., UNNIKRISHNAN, M., KRAUSZ, T. & COHEN, J. (2000) Mitogenic factor (MF) is the major DNase of serotype M89 *Streptococcus pyogenes*. *Microbiology*, 146 (Pt 11), 2785-92.
- STARR, C. R. & ENGLEBERG, N. C. (2006) Role of hyaluronidase in subcutaneous spread and growth of group A streptococcus. *Infect Immun*, 74, 40-8.
- STOLLERMAN, G. H., KANTOR, F. S. & GORDON, B. D. (1958) Accessory plasma factors involved in the bactericidal test for type-specific antibody to group A streptococci. I. A typical behavior of some human and rabbit bloods. *J Exp Med*, 108, 475-91.

- STULBERG, D. L., PENROD, M. A. & BLATNY, R. A. (2002) Common bacterial skin infections. *Am Fam Physician*, 66, 119-24.
- STULKE, J. & HILLEN, W. (2000) Regulation of carbon catabolism in *Bacillus* species. *Annu Rev Microbiol*, 54, 849-80.
- SUMBY, P., BARBIAN, K. D., GARDNER, D. J., WHITNEY, A. R., WELTY, D. M., LONG, R. D., BAILEY, J. R., PARNELL, M. J., HOE, N. P., ADAMS, G. G., DELEO, F. R. & MUSSER, J. M. (2005) Extracellular deoxyribonuclease made by group A *Streptococcus* assists pathogenesis by enhancing evasion of the innate immune response. *Proc Natl Acad Sci U S A*, 102, 1679-84.
- SUMBY, P., WHITNEY, A. R., GRAVISS, E. A., DELEO, F. R. & MUSSER, J. M. (2006) Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog*, 2, e5.
- TALAY, S. R., VALENTIN-WEIGAND, P., JERLSTROM, P. G., TIMMIS, K. N. & CHHATWAL, G. S. (1992) Fibronectin-binding protein of *Streptococcus pyogenes*: sequence of the binding domain involved in adherence of streptococci to epithelial cells. *Infect Immun*, 60, 3837-44.
- TALKINGTON, D. F., SCHWARTZ, B., BLACK, C. M., TODD, J. K., ELLIOTT, J., BREIMAN, R. F. & FACKLAM, R. R. (1993) Association of phenotypic and genotypic characteristics of invasive *Streptococcus pyogenes* isolates with clinical components of streptococcal toxic shock syndrome. *Infect Immun*, 61, 3369-74.
- TENG, F., NANNINI, E. C. & MURRAY, B. E. (2005) Importance of gls24 in virulence and stress response of *Enterococcus faecalis* and use of the GlS24 protein as a possible immunotherapy target. *J Infect Dis*, 191, 472-80.
- TERAO, Y., KAWABATA, S., NAKATA, M., NAKAGAWA, I. & HAMADA, S. (2002) Molecular characterization of a novel fibronectin-binding protein of

Streptococcus pyogenes strains isolated from toxic shock-like syndrome patients. *J Biol Chem*, 277, 47428-35.

TERLECKYJ, B., WILLETT, N. P. & SHOCKMAN, G. D. (1975) Growth of several cariogenic strains of oral streptococci in a chemically defined medium. *Infect Immun*, 11, 649-55.

TETTELIN, H., NELSON, K. E., PAULSEN, I. T., EISEN, J. A., READ, T. D., PETERSON, S., HEIDELBERG, J., DEBOY, R. T., HAFT, D. H., DODSON, R. J., DURKIN, A. S., GWINN, M., KOLONAY, J. F., NELSON, W. C., PETERSON, J. D., UMayAM, L. A., WHITE, O., SALZBERG, S. L., LEWIS, M. R., RADUNE, D., HOLTZAPPLE, E., KHOURI, H., WOLF, A. M., UTTERBACK, T. R., HANSEN, C. L., MCDONALD, L. A., FELDBLYUM, T. V., ANGIUOLI, S., DICKINSON, T., HICKEY, E. K., HOLT, I. E., LOFTUS, B. J., YANG, F., SMITH, H. O., VENTER, J. C., DOUGHERTY, B. A., MORRISON, D. A., HOLLINGSHEAD, S. K. & FRASER, C. M. (2001) Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science*, 293, 498-506.

THIBODEAU, G. A. & PATTON, K. T. (1992) *The human body in health and disease*, London, St. Louis : Mosby Year Book.

THONGBOONKARD, V., LUENGPAILIN, J., CAO, J., PIERCE, W. M., CAI, J., KLEIN, J. B. & DOYLE, R. J. (2002) Fluoride exposure attenuates expression of *Streptococcus pyogenes* virulence factors. *J Biol Chem*, 277, 16599-605.

TITGEMEYER, F. & HILLEN, W. (2002) Global control of sugar metabolism: a gram-positive solution. *Antonie Van Leeuwenhoek*, 82, 59-71.

TYLEWSKA, S., HJERTEN, S. & WADSTROM, T. (1981) Effect of subinhibitory concentrations of antibiotics on the adhesion of *Streptococcus pyogenes* to pharyngeal epithelial cells. *Antimicrob Agents Chemother*, 20, 563-6.

- TYLEWSKA, S. & HRYNIEWICZ, W. (1987) Streptococcus pyogenes cell wall protein responsible for binding to pharyngeal epithelial cells. *Zentralbl Bakteriol Mikrobiol Hyg [A]*, 265, 146-50.
- TYLEWSKA, S., KLOSINSKA-KITA, E. & MALINOWSKI, W. (1978) Chemotactic activity of polymorphonuclear leukocytes to Streptococcus pyogenes. *Zentralbl Bakteriol [Orig A]*, 241, 294-300.
- VAN DE BEEK, D., DE GANS, J., SPANJAARD, L., SELA, S., VERMEULEN, M. & DANKERT, J. (2002) Group A streptococcal meningitis in adults: report of 41 cases and a review of the literature. *Clin Infect Dis*, 34, e32-6.
- VAN DEN HOOGEN, B. M., VAN WEEREN, P. R., LOPES-CARDOZO, M., VAN GOLDE, L. M., BARNEVELD, A. & VAN DE LEST, C. H. (1998) A microtiter plate assay for the determination of uronic acids. *Anal Biochem*, 257, 107-11.
- VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J., SUTTON, G. G., SMITH, H. O., YANDELL, M., EVANS, C. A., HOLT, R. A., GOCAYNE, J. D., AMANATIDES, P., BALLEW, R. M., HUSON, D. H., WORTMAN, J. R., ZHANG, Q., KODIRA, C. D., ZHENG, X. H., CHEN, L., SKUPSKI, M., SUBRAMANIAN, G., THOMAS, P. D., ZHANG, J., GABOR MIKLOS, G. L., NELSON, C., BRODER, S., CLARK, A. G., NADEAU, J., MCKUSICK, V. A., ZINDER, N., LEVINE, A. J., ROBERTS, R. J., SIMON, M., SLAYMAN, C., HUNKAPILLER, M., BOLANOS, R., DELCHER, A., DEW, I., FASULO, D., FLANIGAN, M., FLOREA, L., HALPERN, A., HANNENHALLI, S., KRAVITZ, S., LEVY, S., MOBARRY, C., REINERT, K., REMINGTON, K., ABU-THREIDEH, J., BEASLEY, E., BIDDICK, K., BONAZZI, V., BRANDON, R., CARGILL, M., CHANDRAMOULISWARAN, I., CHARLAB, R., CHATURVEDI, K., DENG, Z., DI FRANCESCO, V., DUNN, P., EILBECK, K., EVANGELISTA, C., GABRIELIAN, A. E., GAN, W., GE, W., GONG, F., GU, Z., GUAN, P., HEIMAN, T. J., HIGGINS, M. E., JI, R. R., KE, Z., KETCHUM, K. A., LAI, Z., LEI, Y., LI, Z., LI, J., LIANG, Y., LIN, X., LU, F., MERKULOV, G. V.,

- MILSHINA, N., MOORE, H. M., NAIK, A. K., NARAYAN, V. A., NEELAM, B., NUSSKERN, D., RUSCH, D. B., SALZBERG, S., SHAO, W., SHUE, B., SUN, J., WANG, Z., WANG, A., WANG, X., WANG, J., WEI, M., WIDES, R., XIAO, C., YAN, C., et al. (2001) The sequence of the human genome. *Science*, 291, 1304-51.
- VIRTANEVA, K., PORCELLA, S. F., GRAHAM, M. R., IRELAND, R. M., JOHNSON, C. A., RICKLEFS, S. M., BABAR, I., PARKINS, L. D., ROMERO, R. A., CORN, G. J., GARDNER, D. J., BAILEY, J. R., PARNELL, M. J. & MUSSER, J. M. (2005) Longitudinal analysis of the group A *Streptococcus* transcriptome in experimental pharyngitis in cynomolgus macaques. *Proc Natl Acad Sci U S A*, 102, 9014-9.
- VOYICH, J. M., STURDEVANT, D. E., BRAUGHTON, K. R., KOBAYASHI, S. D., LEI, B., VIRTANEVA, K., DORWARD, D. W., MUSSER, J. M. & DELEO, F. R. (2003) Genome-wide protective response used by group A *Streptococcus* to evade destruction by human polymorphonuclear leukocytes. *Proc Natl Acad Sci U S A*, 100, 1996-2001.
- WANG, L., FABRET, C., KANAMARU, K., STEPHENSON, K., DARTOIS, V., PEREGO, M. & HOCH, J. A. (2001) Dissection of the functional and structural domains of phosphorelay histidine kinase A of *Bacillus subtilis*. *J Bacteriol*, 183, 2795-802.
- WANNAMAKER, L. W. & YASMINEH, W. (1967) Streptococcal nucleases. I. Further studies on the A, B, and C enzymes. *J Exp Med*, 126, 475-96.
- WARNER, J. B. & LOLKEMA, J. S. (2003) CcpA-dependent carbon catabolite repression in bacteria. *Microbiol Mol Biol Rev*, 67, 475-90.
- WASINGER, V. C., CORDWELL, S. J., CERPA-POLJAK, A., YAN, J. X., GOOLEY, A. A., WILKINS, M. R., DUNCAN, M. W., HARRIS, R., WILLIAMS, K. L. & HUMPHERY-SMITH, I. (1995) Progress with gene-

product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*, 16, 1090-4.

WESSELS, M. R. & BRONZE, M. S. (1994) Critical role of the group A streptococcal capsule in pharyngeal colonization and infection in mice. *Proc Natl Acad Sci U S A*, 91, 12238-42.

WESSELS, M. R., GOLDBERG, J. B., MOSES, A. E. & DICESARE, T. J. (1994) Effects on virulence of mutations in a locus essential for hyaluronic acid capsule expression in group A streptococci. *Infect Immun*, 62, 433-41.

WESSELS, M. R., MOSES, A. E., GOLDBERG, J. B. & DICESARE, T. J. (1991) Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. *Proc Natl Acad Sci U S A*, 88, 8317-21.

WHATMORE, A. M. & KEHOE, M. A. (1994) Horizontal gene transfer in the evolution of group A streptococcal emm-like genes: gene mosaics and variation in Vir regulons. *Mol Microbiol*, 11, 363-74.

WHITNACK, E. & BEACHEY, E. H. (1982) Antiopsonic activity of fibrinogen bound to M protein on the surface of group A streptococci. *J Clin Invest*, 69, 1042-5.

WHITNACK, E., BISNO, A. L. & BEACHEY, E. H. (1981) Hyaluronate capsule prevents attachment of group A streptococci to mouse peritoneal macrophages. *Infect Immun*, 31, 985-91.

WHITNACK, E., DALE, J. B. & BEACHEY, E. H. (1984) Common protective antigens of group A streptococcal M proteins masked by fibrinogen. *J Exp Med*, 159, 1201-12.

WILKINS, J. C., BEIGHTON, D. & HOMER, K. A. (2003) Effect of acidic pH on expression of surface-associated proteins of *Streptococcus oralis*. *Appl Environ Microbiol*, 69, 5290-6.

- WILKINS, J. C., HOMER, K. A. & BEIGHTON, D. (2002) Analysis of *Streptococcus mutans* proteins modulated by culture under acidic conditions. *Appl Environ Microbiol*, 68, 2382-90.
- WILM, M. & MANN, M. (1996) Analytical properties of the nanoelectrospray ion source. *Anal Chem*, 68, 1-8.
- YAN, J. X., DEVENISH, A. T., WAIT, R., STONE, T., LEWIS, S. & FOWLER, S. (2002) Fluorescence two-dimensional difference gel electrophoresis and mass spectrometry based proteomic analysis of *Escherichia coli*. *Proteomics*, 2, 1682-98.

Appendices

Appendix A

A1: Chemicals, media and enzymes used in this study.

Acros

Adenine
p-aminobenzoate
Ammonia solution
Cysteine
3, 5-Dinitrosalicylic acid (DNSA)
Disodium phosphate
Guanine
Iodoacetamide
Monosodium phosphate
2,7 – naphthalene-disulfonic acid solution
Nicotinamide
Sodium hyaluronate
Uracil

Aldrich

Pantothenate

British Drugs House (Merck)

Biotin
Folic acid

Fisher BioReagents

Acrylamide/Bisacrylamide 37.5:1, 40 % solution
Brilliant blue G 250
Coomassie Blue R-250
Lysozyme
Methonal
Phenol

Fisher Chemicals

Acetic acid, glacial
Aceton
Acetonitrile
Ammonium bicarbonate
Ammonium sulphate

Dimethyl sulfoxide
Ethanol
Formaldehyde
Glutaraldehyde
Phosphoric acid
Sodium carbonate
Sodium tetraborate
Trifluoroacetic acid
Trisodium citrate

Melford Laboratories Ltd

Agarose (High gel strength)
Dithiothreitol
Glycine
Sodium dodecylsulphate
Tris [Hydroxymethyl] aminomethane (Tris-HCl)

Promega

Trypsin Gold

Riedel-deHaen

Hydrochloric acid
Sulfuric acid

Sigma

Albumin
Ammonium sulphate
Angiotensin I
APS
Bovine serum albumin, Fraction V (BSA)
Bromophenol blue
Carbonic anhydrase
CHAPS
Citric acid
Dipotassium phosphate
Ethidium bromide
Ethylene diamine tetraacetic acid, disodium salt (EDTA)
Ferrous sulphate
 β -Galactosidase
Glucose
Glucuronic acid
L-glutamic dehydrogenase
Glyceraldehydes-3-phosphate dehydrogenase
Glycerol
Isopropanol (2- propanol)
 α - lactalbumin

Manganous sulphate
Magnesium chloride
 β -mercaptoethanol
Ovalbumin
Phosphorylase b
Potassium Chloride
Potassium phosphate monobasic
Riboflavin
Silver nitrate
Sodium acetate
Sodium chloride
Sodium hydroxide
Sodium sulphite
Sucrose
TEMED
Thiamine
Trypsinogen
Urea

A2: Protein size standards

Sigma:

Proteins	M.W. (Da)	High	Low
Myosin, rabbit muscle	205000	X	
β -Galactosidase, <i>E. coli</i>	116000	X	
Phosphorylase b, rabbit muscle	97000	X	
Fructose-6-phosphate Kinase, rabbit muscle	84000	X	
Albumin, bovine serum	66000	X	X
Glutamic Dehydrogenase, bovine liver	55000	X	
Ovalbumin, chicken egg	45000	X	X
Glyceraldehydes-3-phosphate Dehydrogenase, rabbit muscle	36000	X	X
Carbonic Anhydrase, bovine erythrocytes	29000		X
Trypsinogen, bovine pancreas	24000		X
Trypsin Inhibitor, soybean	20000		X
α -Lactalbumin, bovine milk	14200		X
Aprotinin, bovine lung	6500		X

Table A1: Molecular weight distribution in Sigma Markers.

A3: 2-D Clean-Up kit

Amersham Biosciences:

The recipes for each solution are not known because the manufacture does not provide any information.

Precipitant solution

This solution renders proteins insoluble.

Co-precipitant solution

This solution contains reagents that co-precipitate with proteins and enhances their removal from solution.

Wash buffer

This solution is used to remove non-protein contaminants from the protein precipitate.

Wash additive

This solution contains a reagent that promotes rapid and complete resuspension of the sample proteins

A4: 2-D quant kit

Amersham Biosciences:

The recipes for each solution are not known because the manufacture does not provide any information.

Precipitant solution

This solution renders proteins insoluble.

Co-precipitant solution

This solution contains reagents that co-precipitate with proteins and enhances their removal from solution.

Copper solution

Precipitated protein is resuspended with this solution.

Colour reagent A

This solution is mixed with colour reagent B to prepare the colour reagent used to measure unbound copper ion.

Colour reagent B

This solution is mixed with colour reagents A to prepare the colour reagent used to measure unbound copper ion.

Bovine serum albumin standard solution

This solution is used to prepare a standard curve.

Appendix B

B1: Autoclaving

Autoclave sterilisation was achieved using a benchtop Prestige[®] Medical 2100 Classic autoclave at 121 °C, 32 lb/inch² pressure for 20 min.

B2: pH meter

All adjustments to the pH of solutions and media were achieved using a Jenway Ion Meter 3340 calibrated with buffers at pH 4.0, 7.0, and 9.2.

B3: Spectrophotometer

Spectrophotometric measurements were made using a UV visible Heλios α Spectronic Unicam spectrophotometer.

B4: Incubators

For the growth of *E. coli* TOP10 in liquid cultures a GallenKamp orbital shaker was used. Growth of bacteria on solid media and growth of GAS in liquid culture were performed in a static GallenKamp incubator.

B5: Centrifugation

Centrifugation of volumes below 1.5ml was achieved using a small Sigma 1-15 bench top micro-centrifuge. Volumes above 1.5ml were centrifuged in a large Sigma 3K18C refrigerated bench top centrifuge, using the rotors and inserts appropriate to the application

Rotor	Vessel containing solution	Insert	RCF	Application
11133	200ml Plastic centrifuge vessel (part N° 15202)	None	4000	Harvest the cells
12131	1.5ml micro-centrifuge tube	None	14000	Extract of protein from lysed cells

Table B1: Centrifuge rotors, vessels and applications.

B6: Sonication

Cell lysis was achieved using a MSE Soniprep 150 ultra-sonication machine.

B7: Microtitre plate reader

All DNSA assays were measured A_{570} using a FL600 FAK microtitre plate reader (Bio-Tek Kontron Instruments Inc.) using MicroTek OS software (MTK version 2.5).

Uronic acid assay was measured A_{570} using an EL808 Ultra microplate reader (BIO-TEK Instruments Inc).

B8: SDS-PAGE gel kit

Electrophoresis of proteins was achieved using a Bio-Rad Mini-PROTEAN® 3 Cell kit powered by an E-C 570-90 E-C apparatus corporation power pack.

B9: Gel documentation

Visualisation of SDS-PAGE gels was achieved using either a CCD camera with Bio-Rad Gel Doc 2000 system or a Bio-Rad GS-710 densitometer system and Quantity One™ software. Hard copies of the gel picture were produced using a Mitsubishi

Video Copy Processor (Model P91), with Mitsubishi thermal paper (K65HM-CE/High density type, 110 mm x 21 m).

B10: Immobiline DryStrip Kit

This kit is designed for running the first dimension of 2-D electrophoresis using precast Immobiline™ DryStrip gels

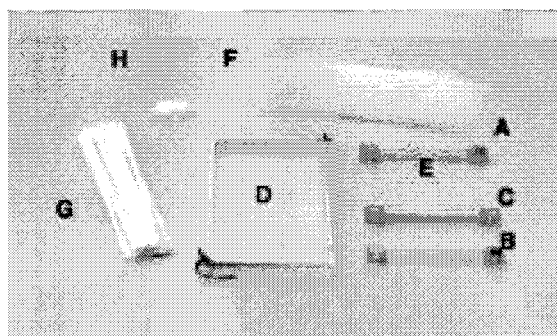


Figure B1 Immobiline DryStrip Kit

http://www5.amershambiosciences.com/aptrix/upp01077.nsf/Content/uk_homepage

- A. Electrode, anode
- B. Electrode, cathode
- C. Sample cup bar
- D. Tray and electrode holder
- E. Sample cups
- F. DryStrip aligners
- G. IEF electrode strips
- H. IEF sample application pieces

B11: Protean II XL 2-D cells and accessories

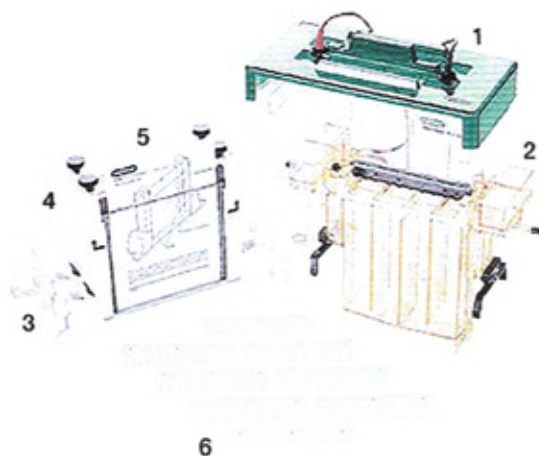


Figure B2 Protean II XL 2-D cells

[http://www.bio-](http://www.bio-rad.com/B2B/BioRad/br_community_home.jsp?BV_SessionID=@@@0449509414.1155220750@@@&BV_EngineID=ccccaddiifdjlihcfcgkmdhkkdfl.0&loggedIn=false&country=null&lang=English&divName=Life+Science+Research)

[rad.com/B2B/BioRad/br_community_home.jsp?BV_SessionID=@@@0449509414.1155220750@@@&BV_EngineID=ccccaddiifdjlihcfcgkmdhkkdfl.0&loggedIn=false&country=null&lang=English&divName=Life+Science+Research](http://www.bio-rad.com/B2B/BioRad/br_community_home.jsp?BV_SessionID=@@@0449509414.1155220750@@@&BV_EngineID=ccccaddiifdjlihcfcgkmdhkkdfl.0&loggedIn=false&country=null&lang=English&divName=Life+Science+Research)

1. Tank and lid
2. Central cooling core
3. Casting stand
4. Sandwich clamps
5. Alignment card
6. Combs

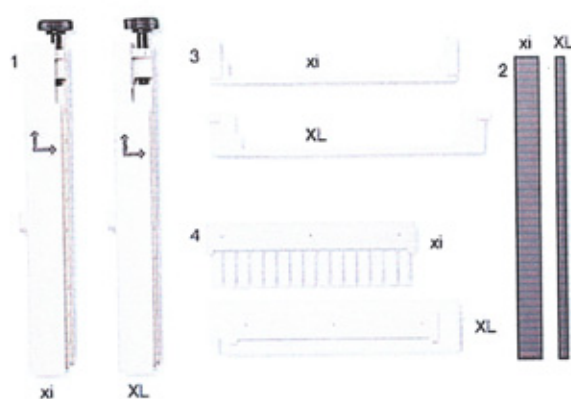


Figure B3 Protean II XL 2-D accessories

1. 4 mm xi clamp notch vs. 13 mm XL clamp notch
2. 19 mm xi spacer vs. 8 mm XL spacer
3. 181 mm xi core gasket vs. 198 mm XL core gasket
4. 153 mm xi comb vs. 184 mm XL comb

B12: Centrifugal evaporator

Extraction of protein digests were dried using CHRIST centrifugal evaporator RVC 2-18

Appendix C Preparation of dialysed sodium hyaluronate

Dialysis of proteins or substrates was achieved in cellulose membrane tubing (size 25 mm x 16 mm; retains > 90 % Cytochrome C (M_r 12 400 Da); Sigma) and secured with dialysis tubing closures (gripping length 50 mm; Sigma).

Purified *CtCBM6* wild type and mutant proteins were dialysed against 3 x 8 L:

Dialysis buffers

Per Litre

NaH₂PO₄ 6.9 g

pH 7.0

Sodium hyaluronate, soluble oat spelt and birchwood xylans were dialysed against 18.2 MΩ/cm H₂O (3 x 4 L).

Preparation of dialysis tubing

The dialysis tubing was cut into pieces of convenient length and boiled for 10 min in a large volume of 18.2 MΩ/cm H₂O. From this point onward, the tubing was always handled with gloves. The tubing was then rinsed thoroughly inside and out with 18.2 MΩ/cm H₂O and once cool, was then stored at 4 °C in 18.2 MΩ/cm H₂O.

Sodium hyaluronate was prepared in the presence of the metal ion chelator EDTA. Sodium hyaluronate (1 g) was added to 100 ml 18.2 MΩ/cm H₂O and autoclaved. This solution was then stirred for 5 h at room temperature with 50 mM EDTA, pH 8.0. Precipitates were then removed by centrifugation for 15 min at 14 000 x g and supernatants were subsequently dialysed three times in 4 L 18.2 MΩ/cm H₂O overnight. The dialysed substrate was then aliquotted in 20 ml volumes into Petri

dishes, frozen at $-20\text{ }^{\circ}\text{C}$, lyophilised overnight and stored at room temperature for future use. Prior to kinetic analysis, the substrate was resuspended in $18.2\text{ M}\Omega/\text{cm H}_2\text{O}$.

Appendix D

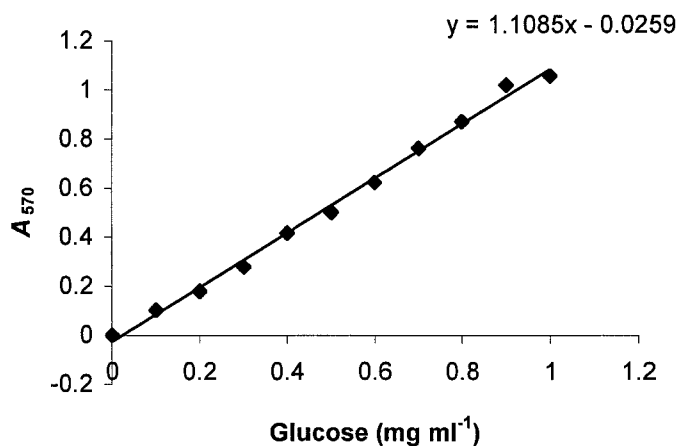


Figure D1. Typical standard curve for DNSA assays.

The curve was generated by the addition of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, + 1.0 mg/ml glucose (final concentration), 1 mg/ml BSA (final concentration) and 20 mM (final concentration) of Tris-HCl pH6.5, in a total volume of 150 μ l. DNSA reagent (150 μ l) was added, the tubes boiled for 20 min and placed on ice for 10 min. The tubes were finally equilibrated to room temperature before the A_{570nm} was determined.

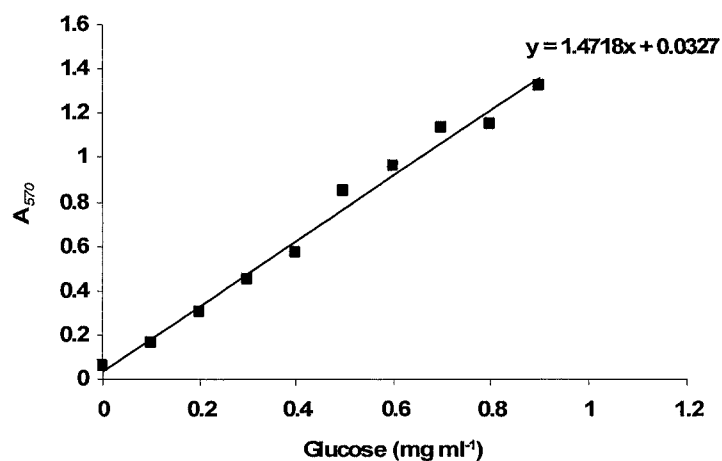


Figure D2. Standard curve for DNSA reducing sugar assays.

The curve was generated by the addition of 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, + 1.0 mg/ml glucose (final concentration) to 18.2 MΩ/cm H₂O in a total volume of 150 μl. DNSA reagent (150 μl) was added, the tubes boiled for 20 min and placed on ice for 10 min. The tubes were finally equilibrated to room temperature before the A₅₇₀ was determined.

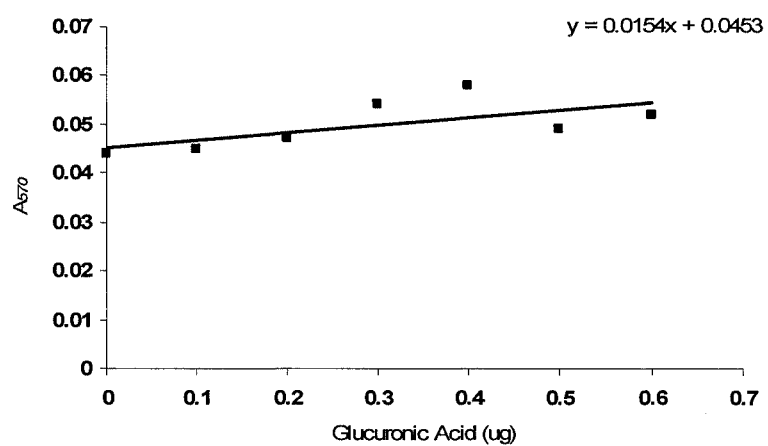


Figure D3. Standard curve for the uronic acid assay

The curve was generated by measure the absorbancies of 0, 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6 μg glucuronic acid at 570 nm using an EL808 Ultra microplate reader.

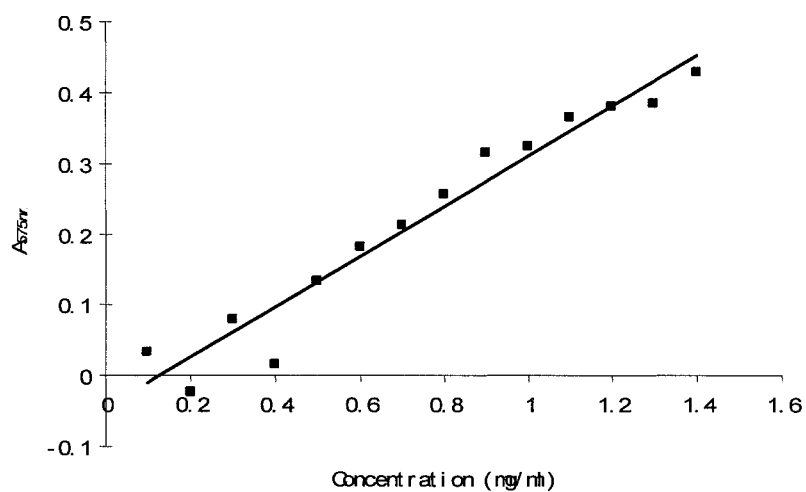


Figure D4. Bradford assay standard curve for the quantification of protein.

The curve was generated by the addition of 5 μ l of 0.1 – 1.4 mg/ml BSA to 250 μ l Bradford's reagent. The absorbance values were measured at 570 nm. $y = 0.3555x - 0.04618$

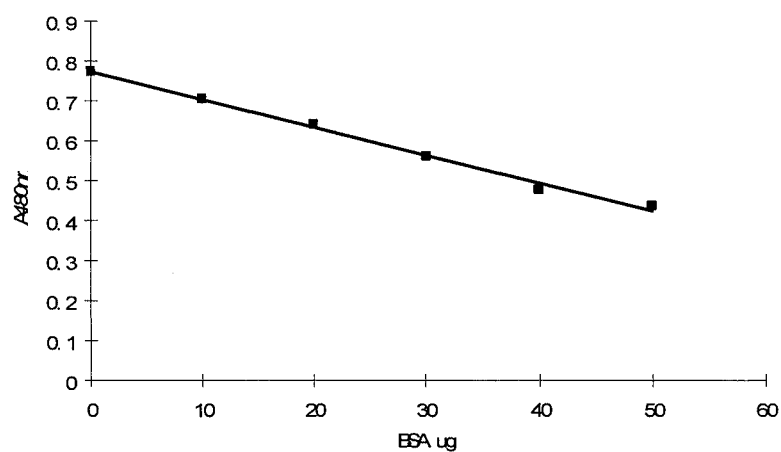
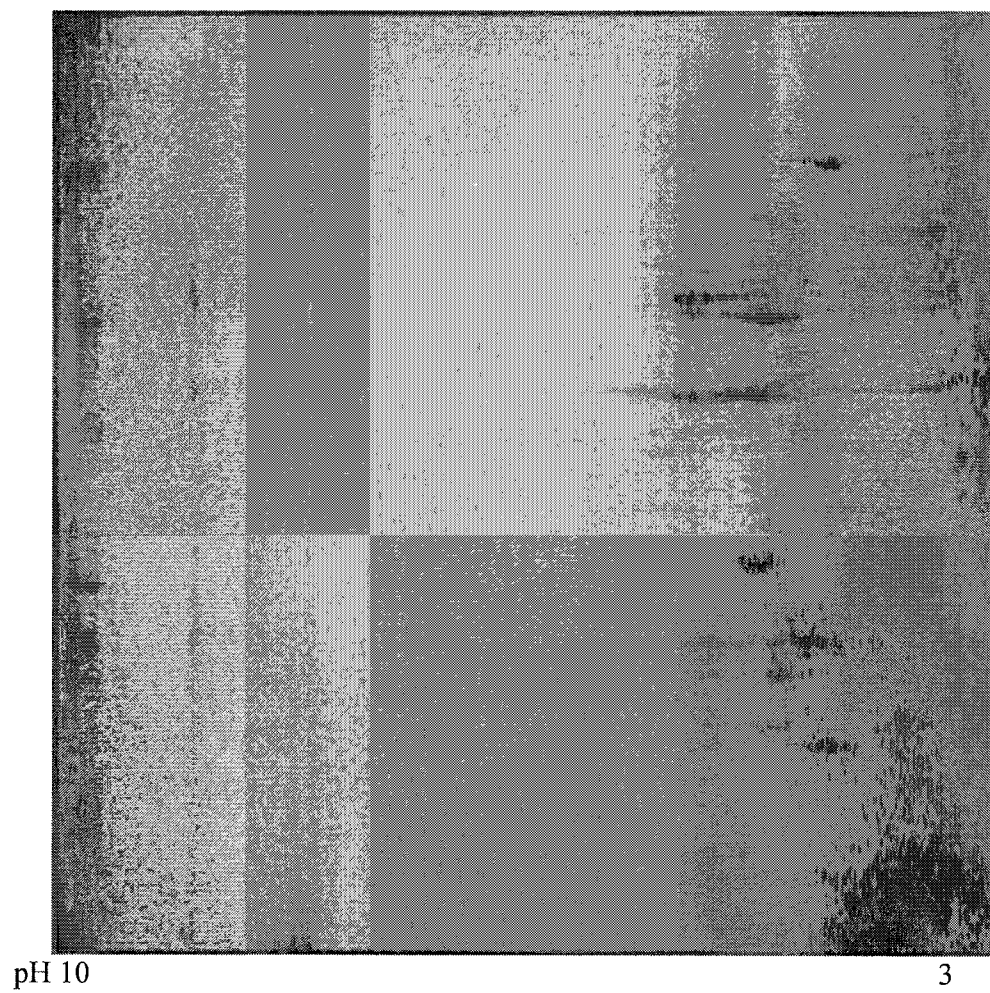


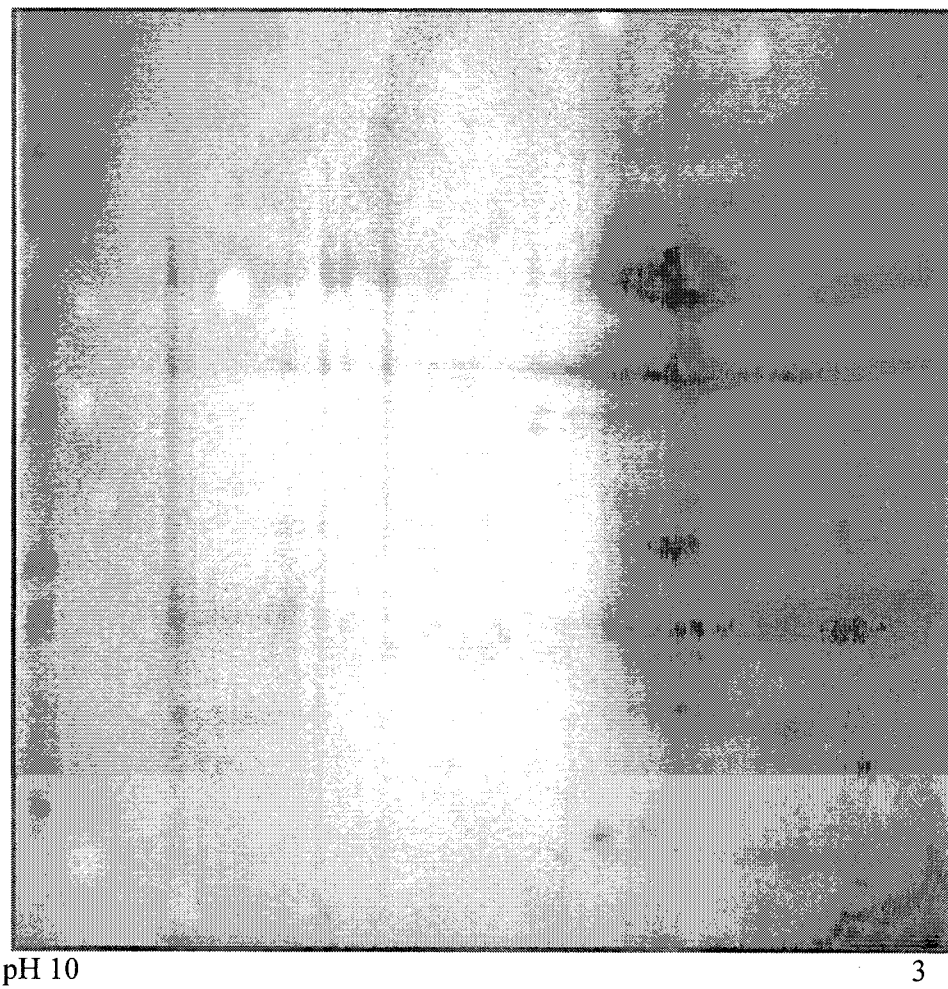
Figure D5. Typical standard curve for assay using 2-D Quant Kit.

The curve was generated by measuring the absorbance values of BSA (0 μg , 10 μg , 20 μg , 30 μg , 40 μg and 50 μg) which were treated with the 2-D Quant Kit at 480 nm.

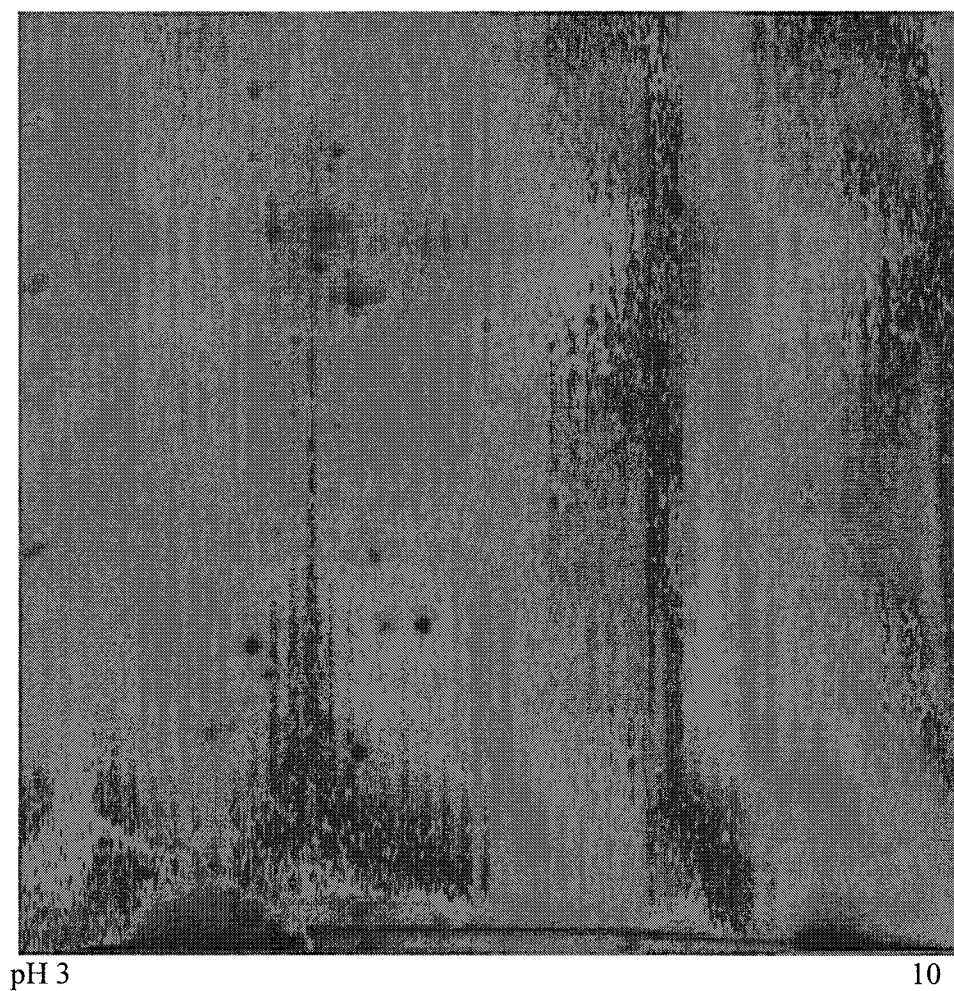
Appendix E



(a) Replicate of 2D gel in experiment No.5



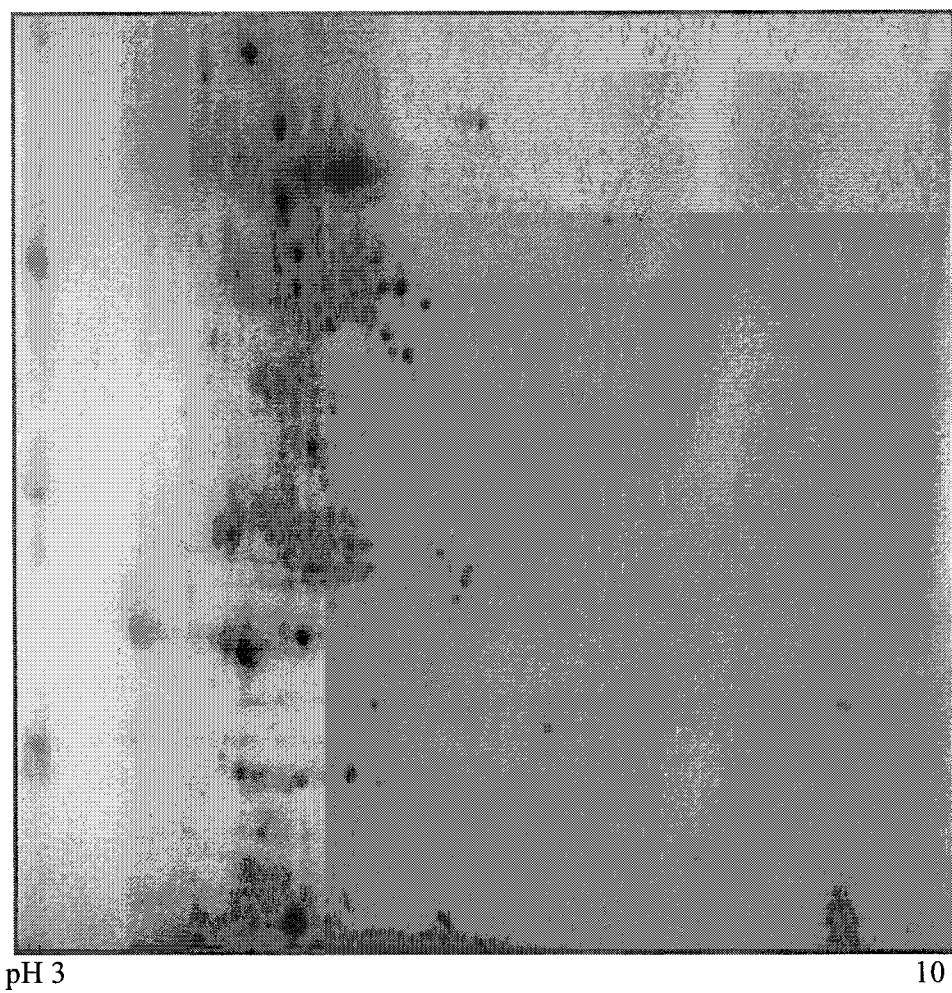
(b) Replicate of 2D gel in experiment No.6



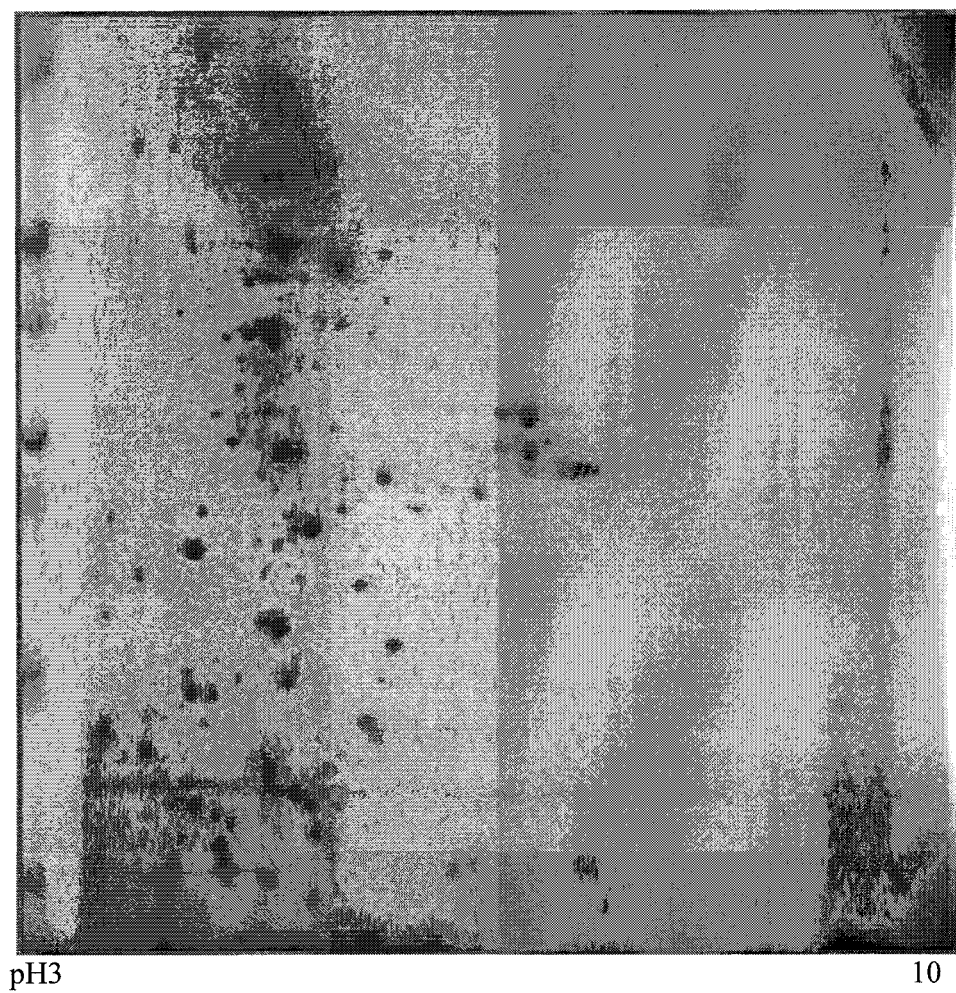
pH 3

10

(c) Replicate of 2D gel in experiment No.7

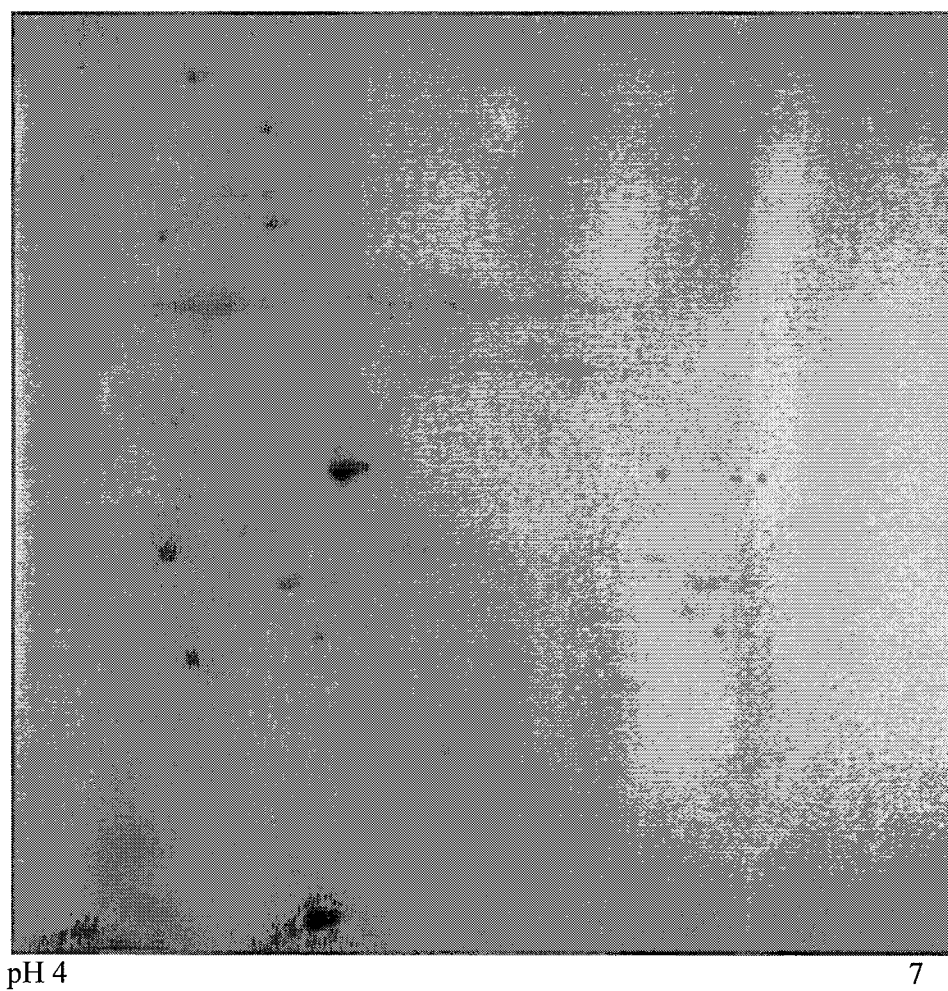


(d) Replicate of 2D gel in experiment No.8

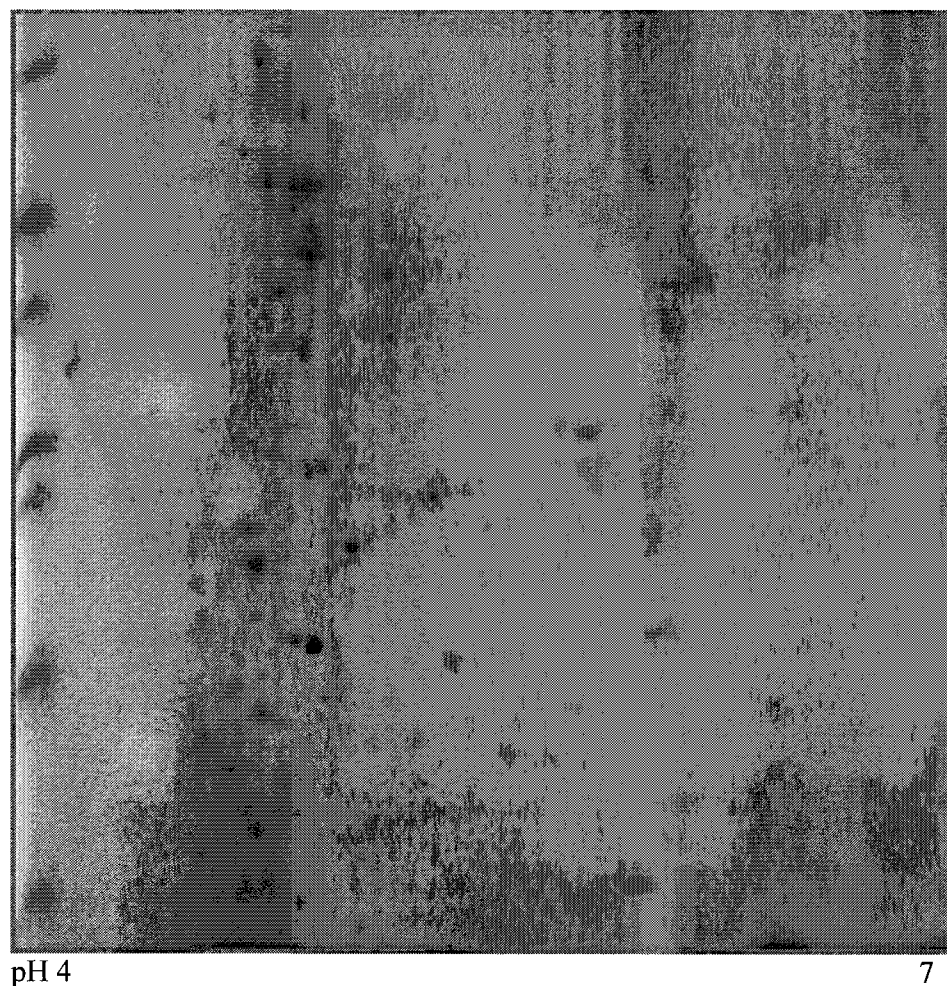


(e) Replicate of 2D gel in experiment No.9

Figure E1. The replicates of the 2D gels in experiment No.5 – 9.



(a) Replicate of 2D gel in experiment No.10



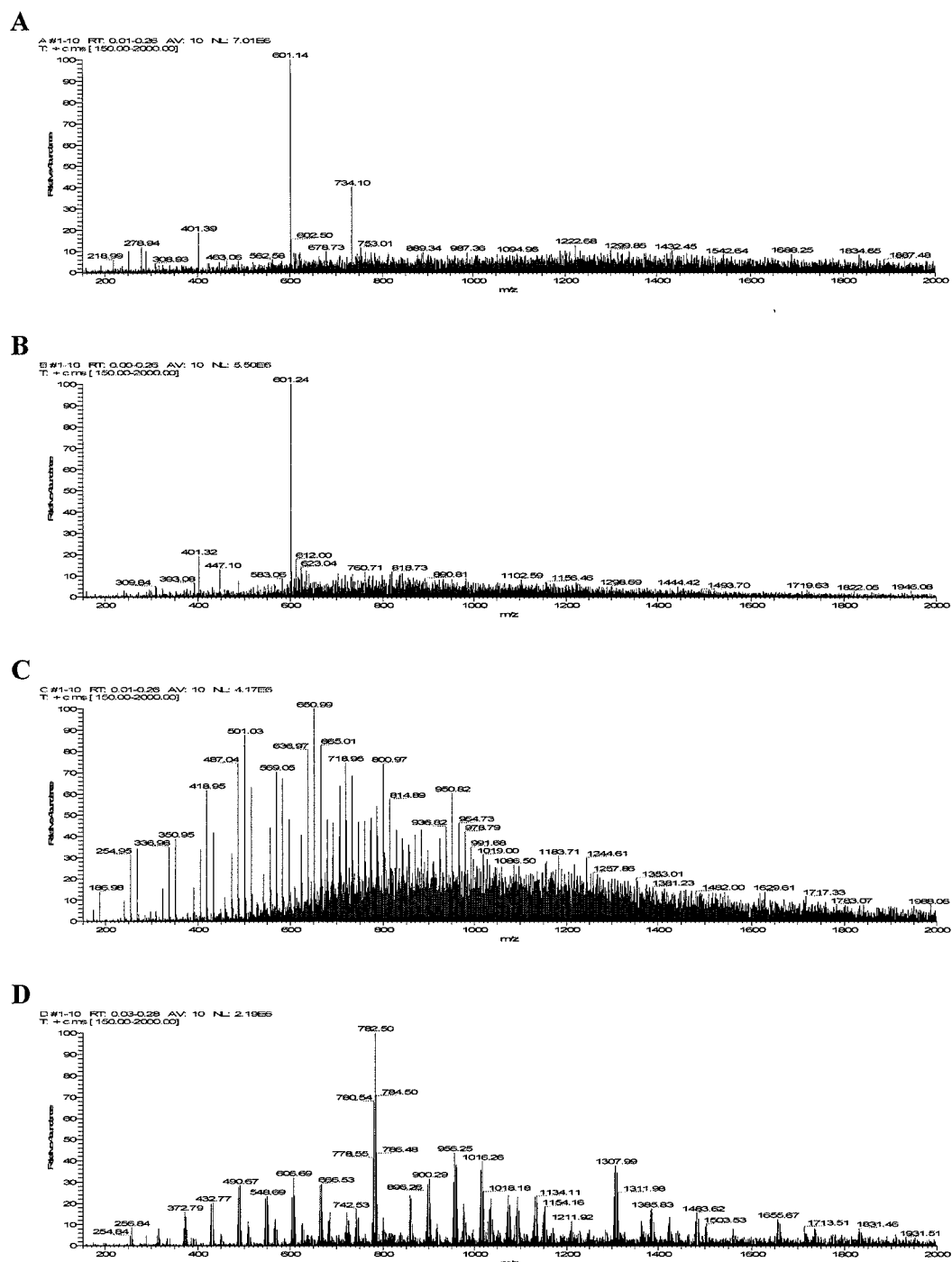
pH 4

7

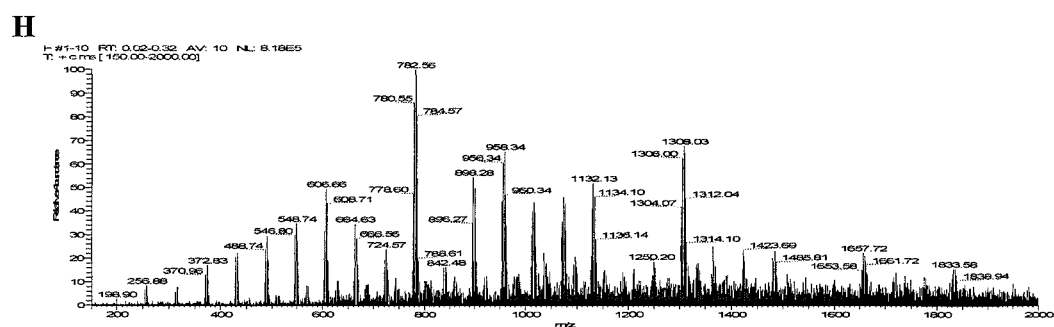
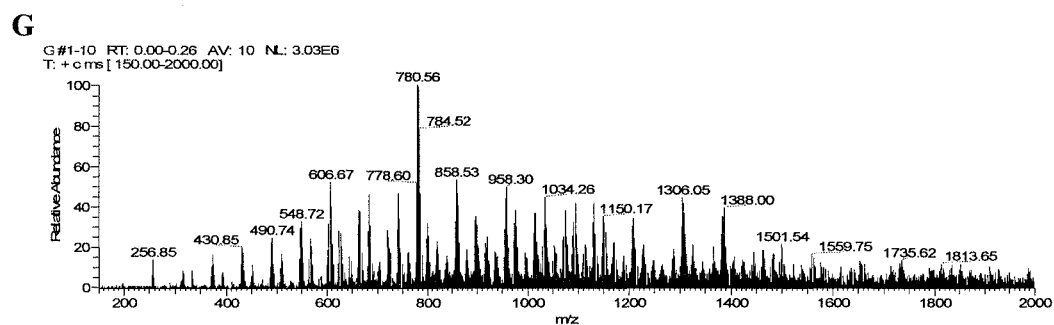
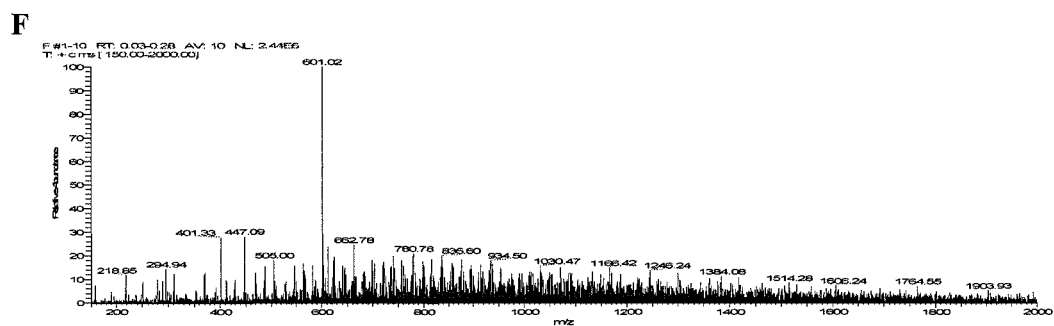
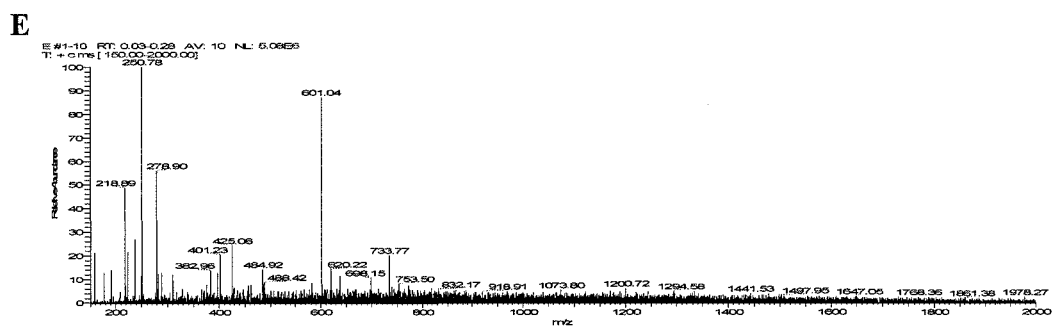
(b) Replicate of 2D gel in experiment No.11

Figure E2. The replicates of the 2D gels in experiment No.10 – 11.

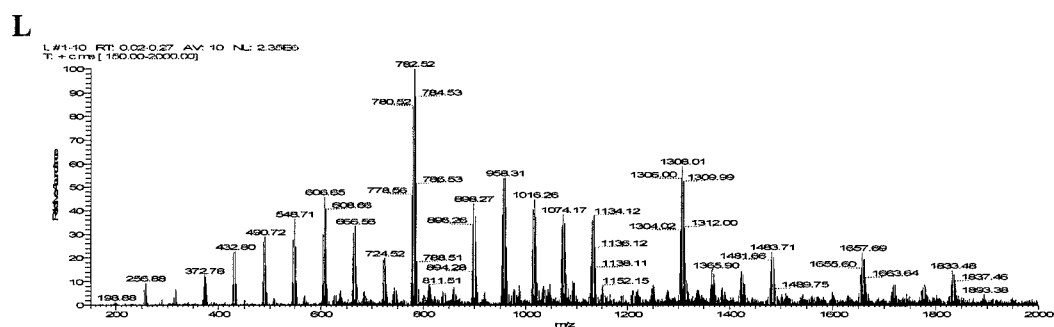
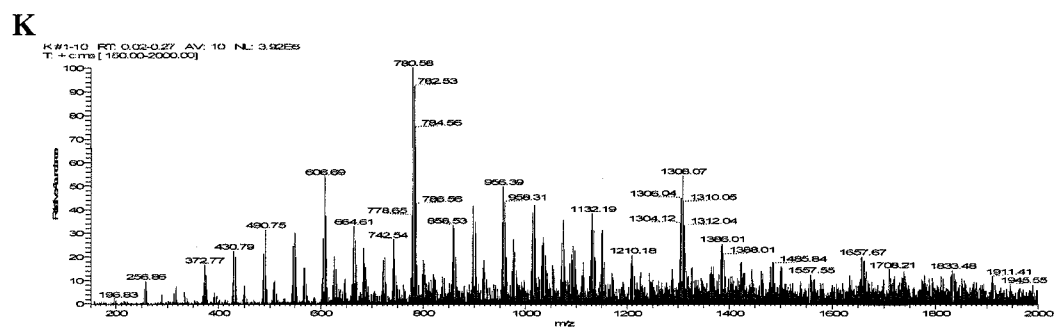
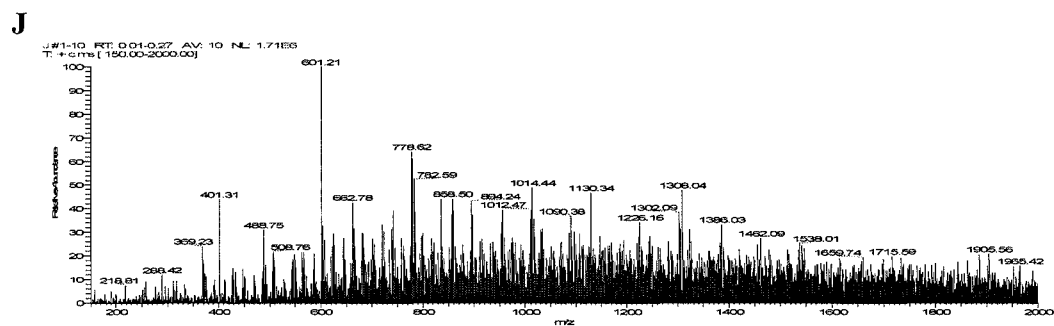
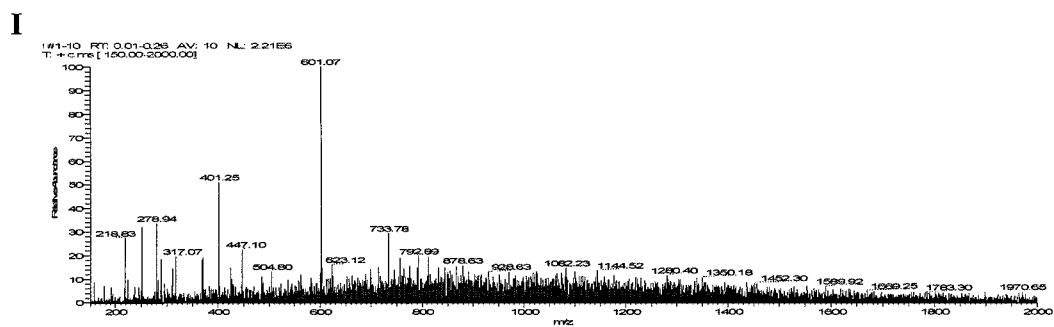
Appendix F



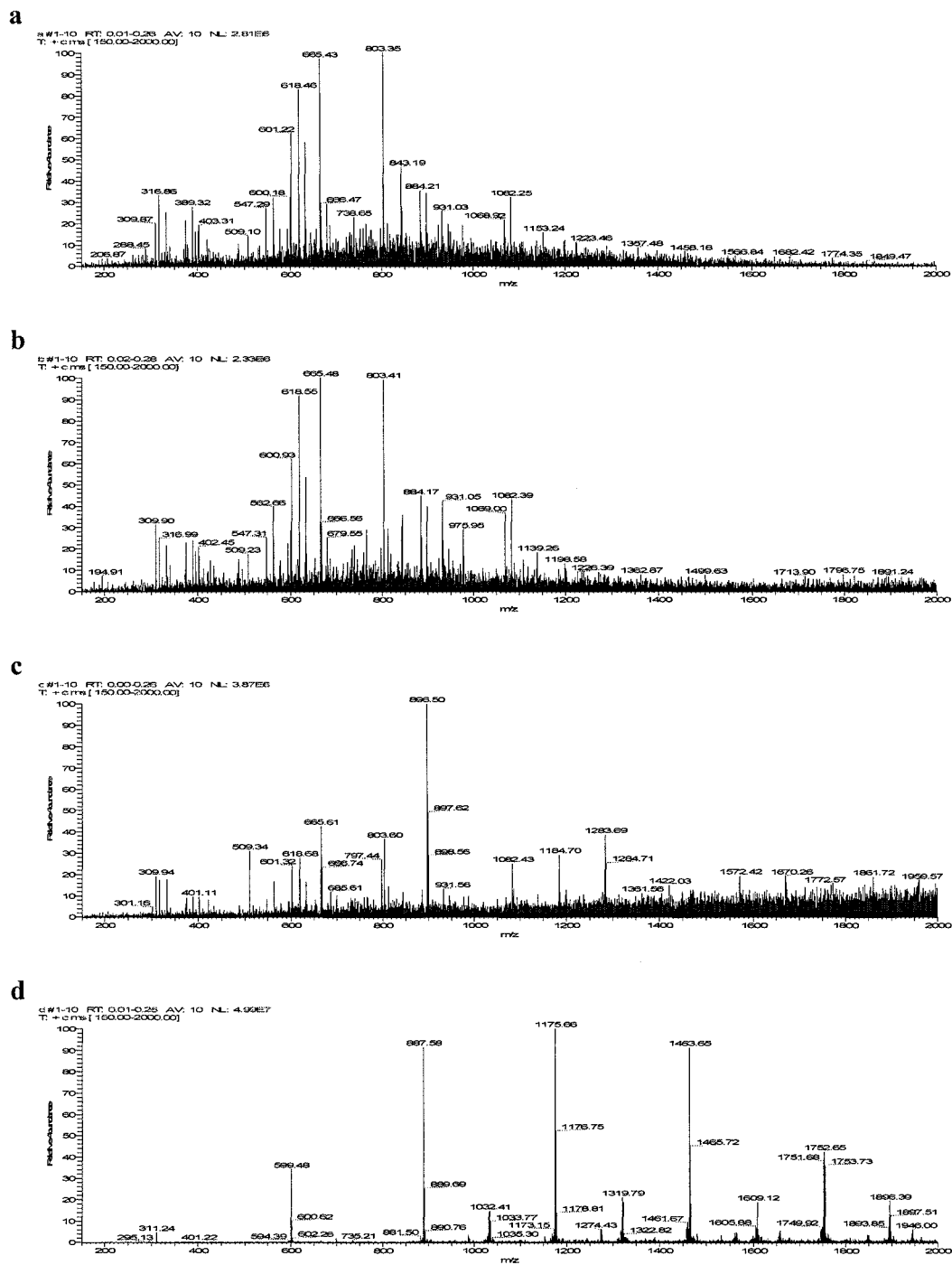
(Figure continues on next page)



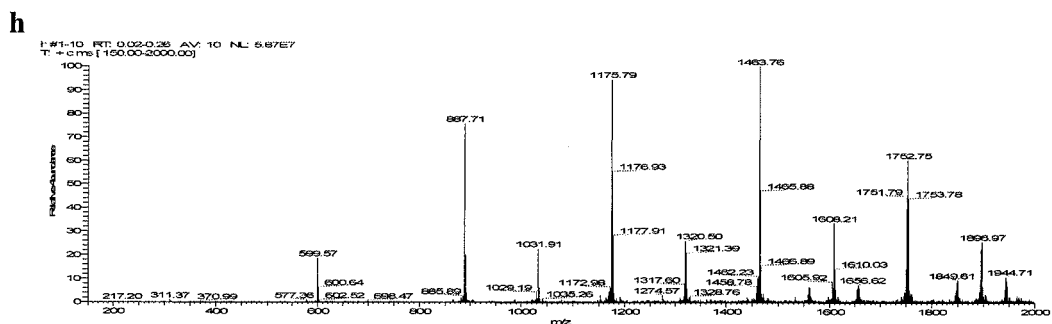
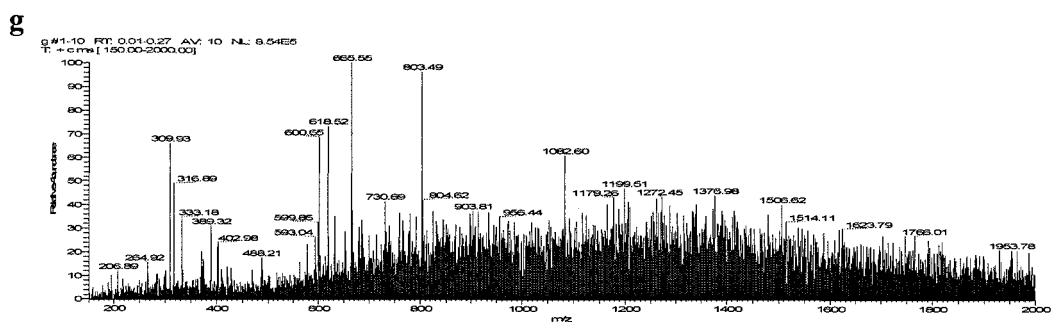
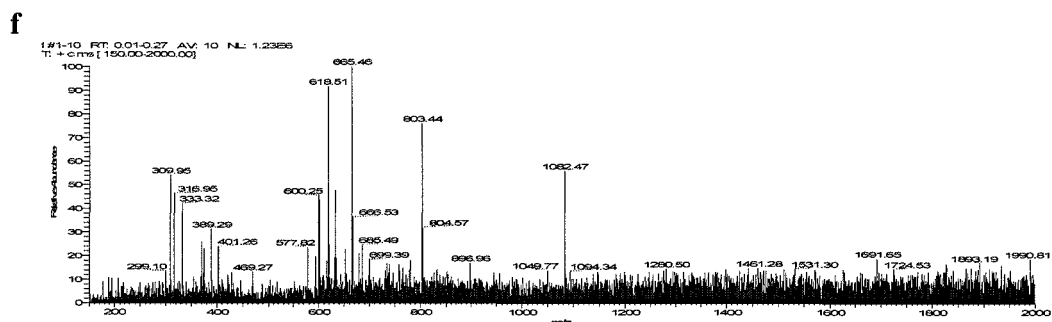
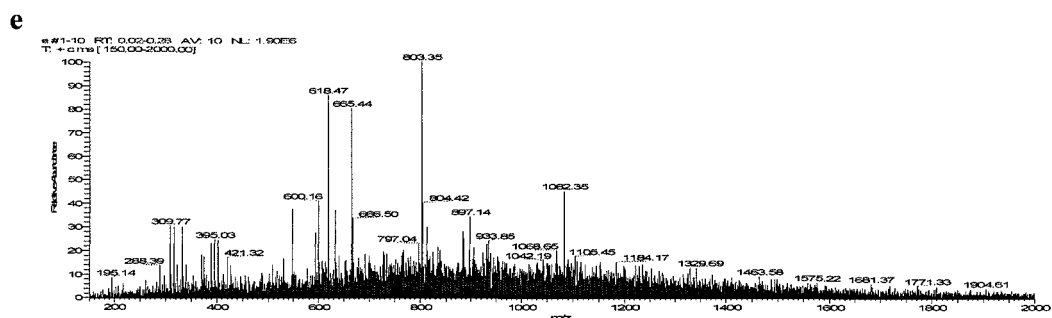
(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)

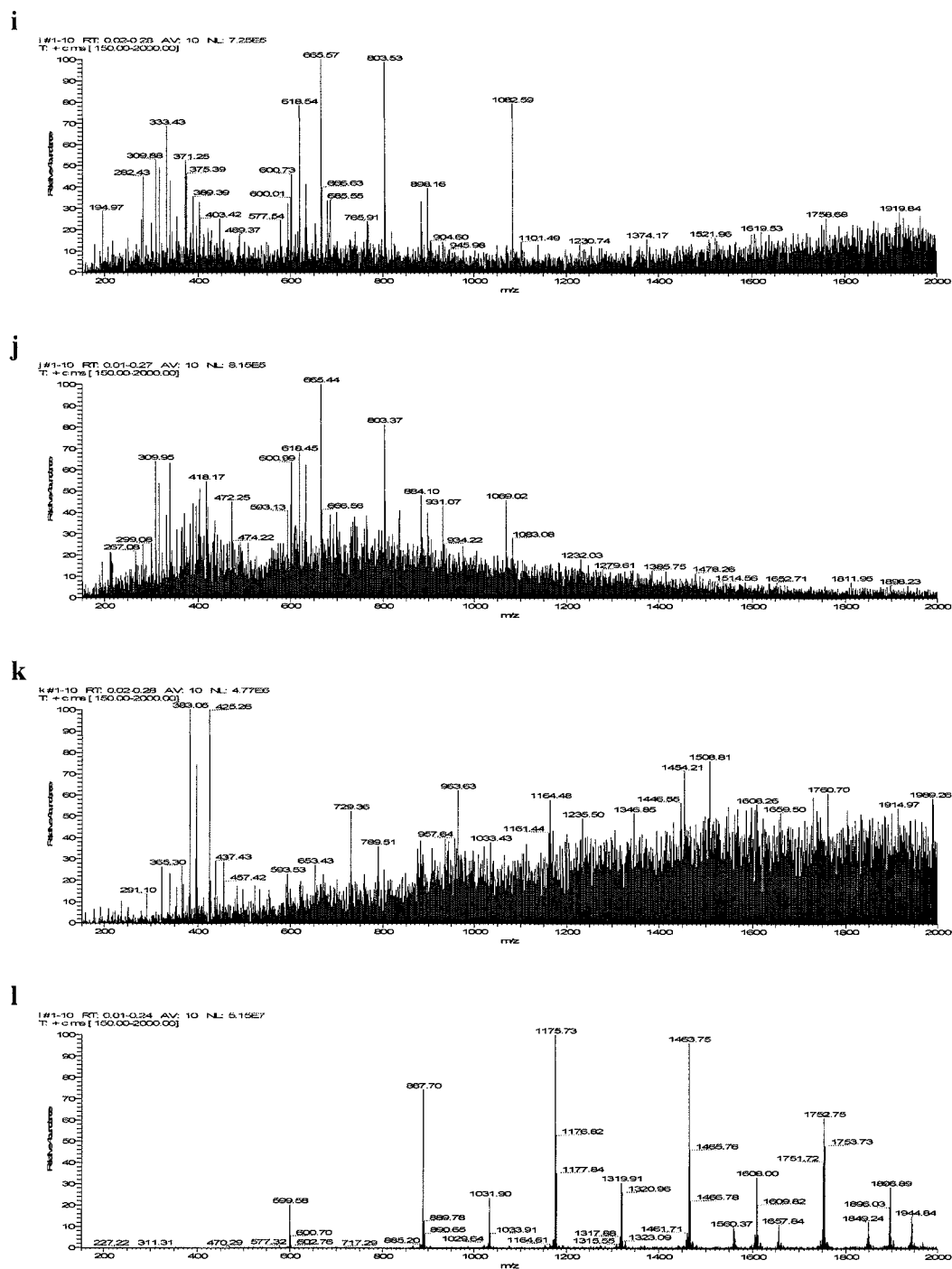
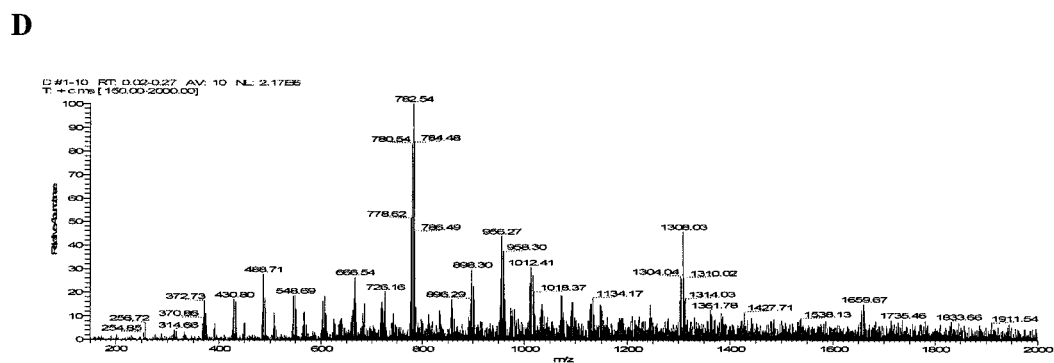
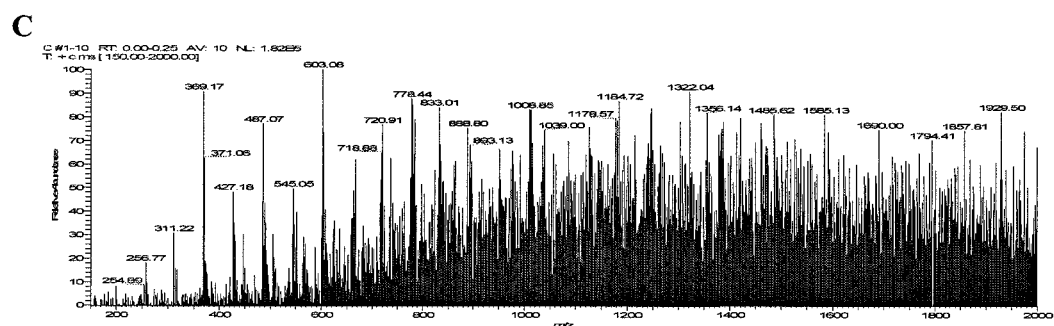
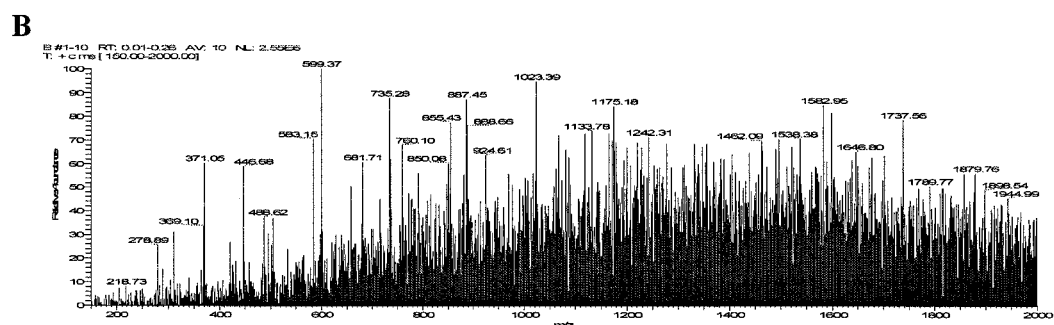
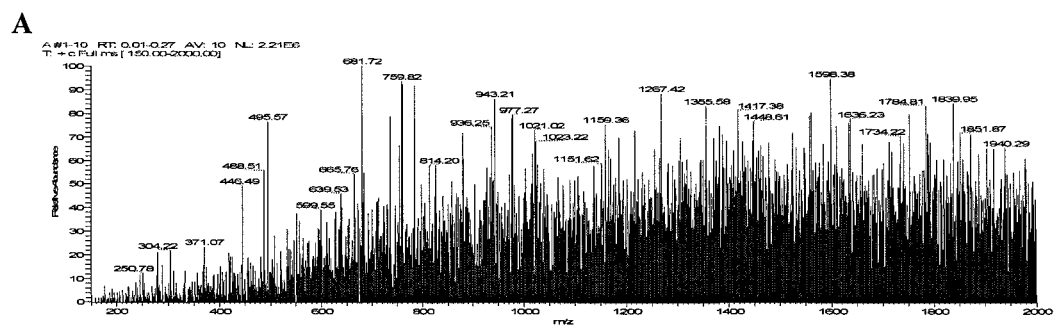
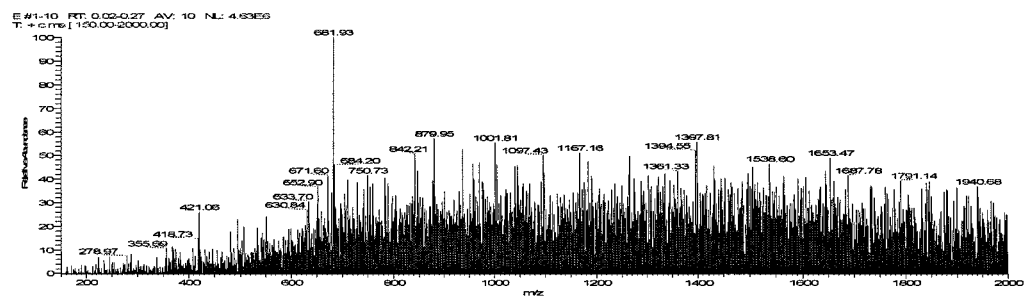
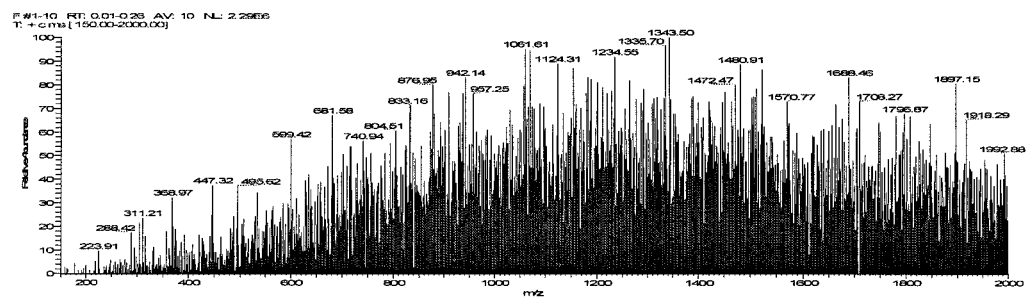
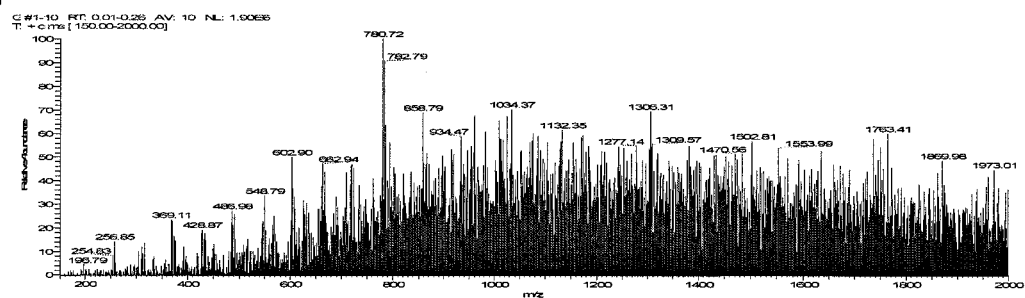
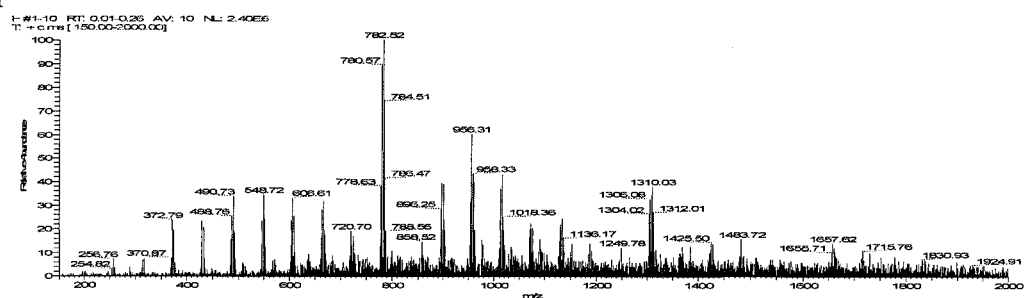


Figure F1 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using α -lactalbumin digests.

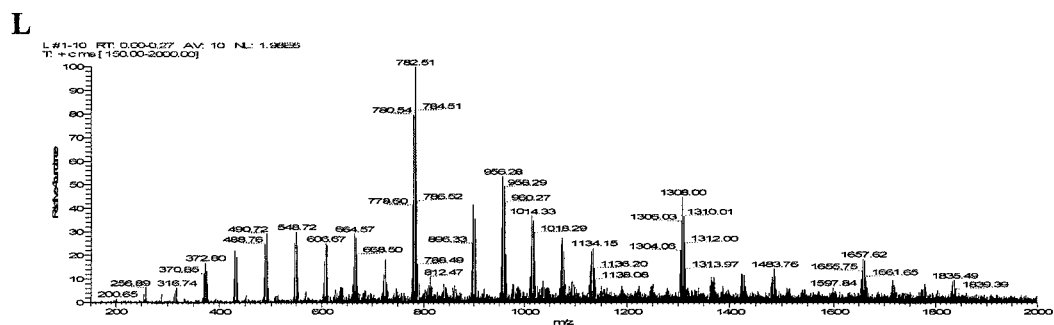
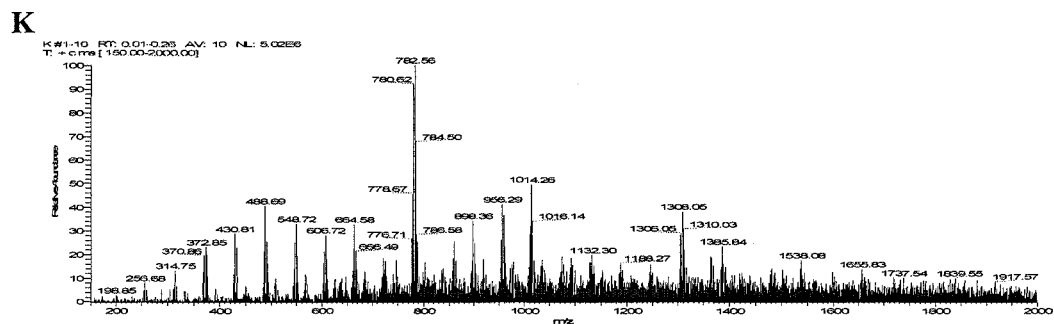
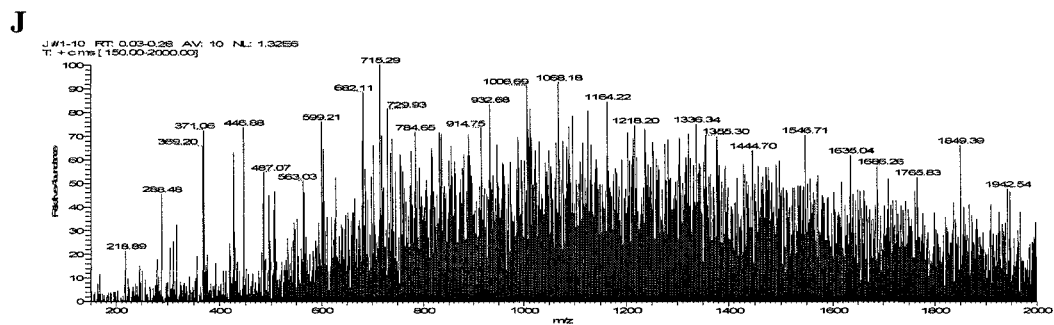
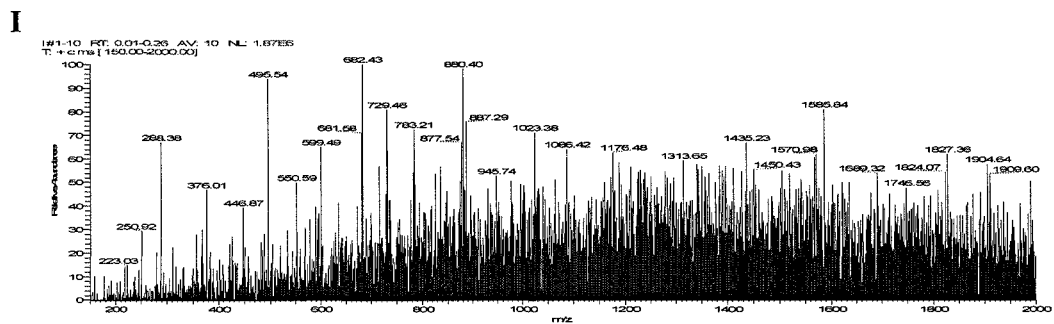
The sample conditions (A-L and a-l) are given in Table 2.10.



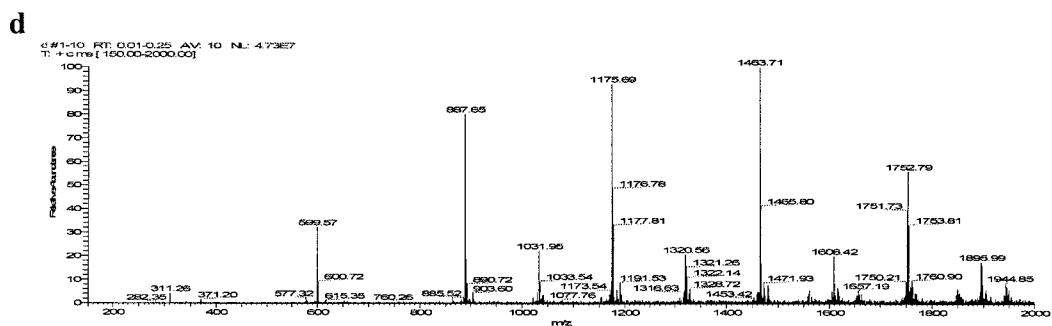
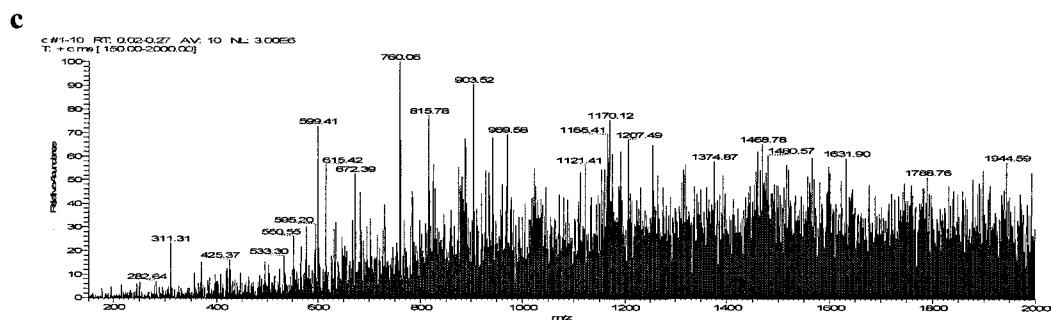
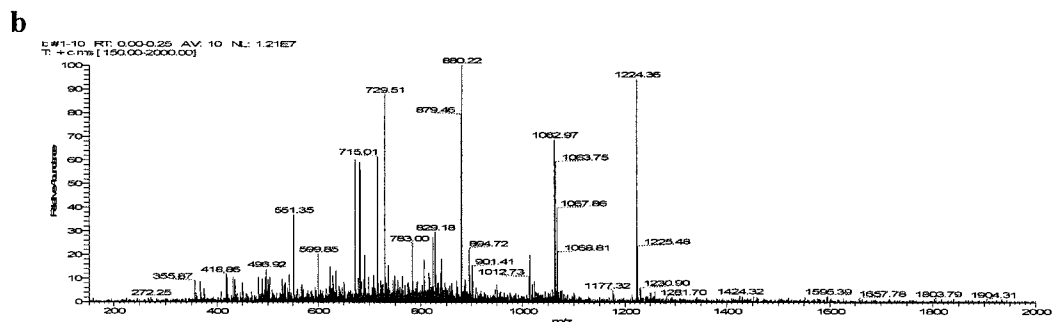
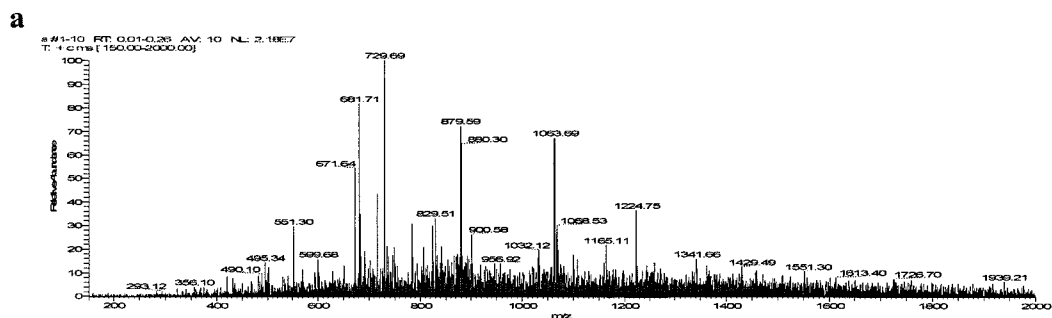
(Figure continues on next page)

E**F****G****H**

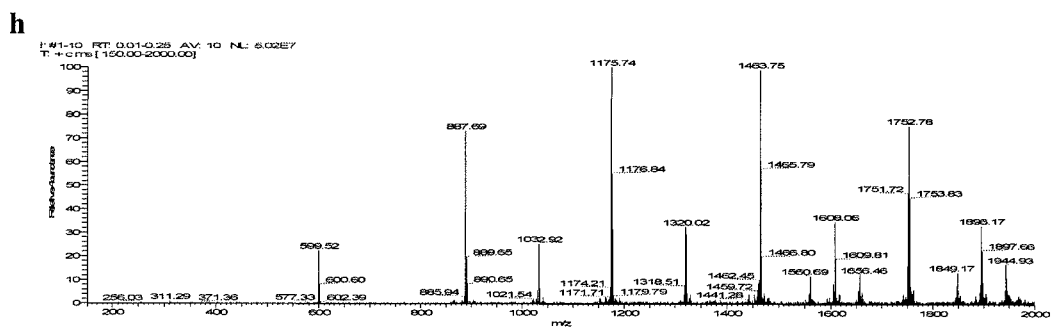
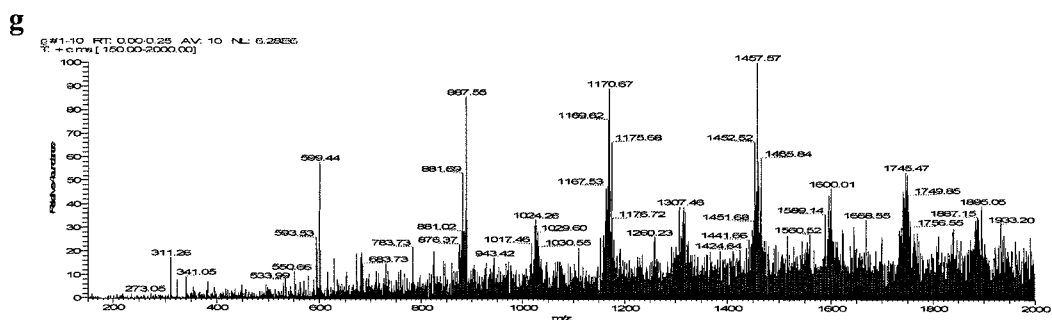
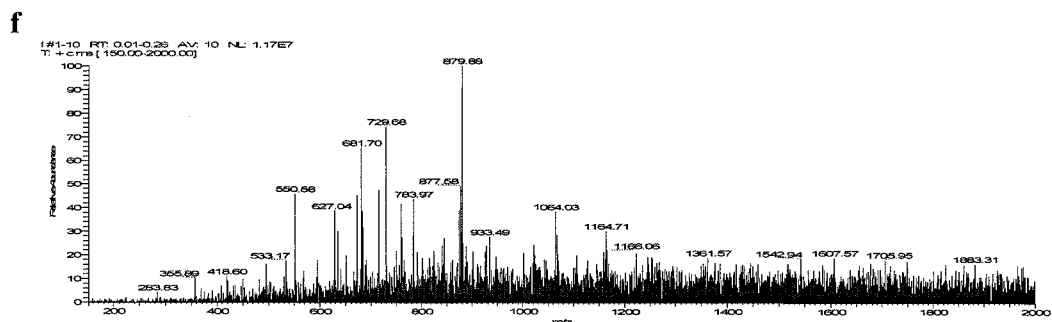
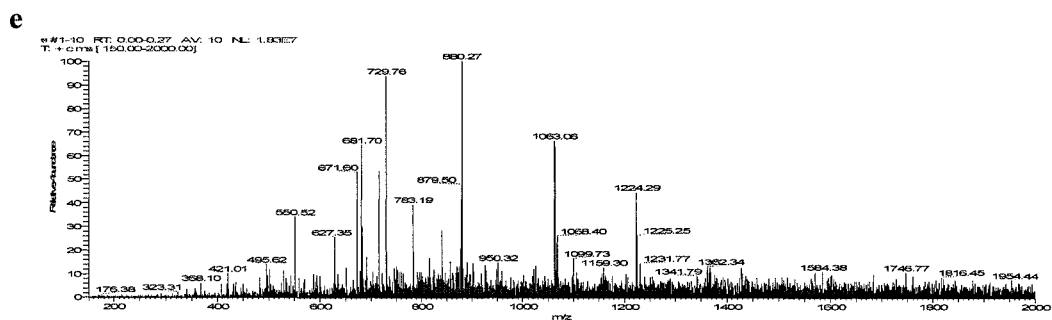
(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)

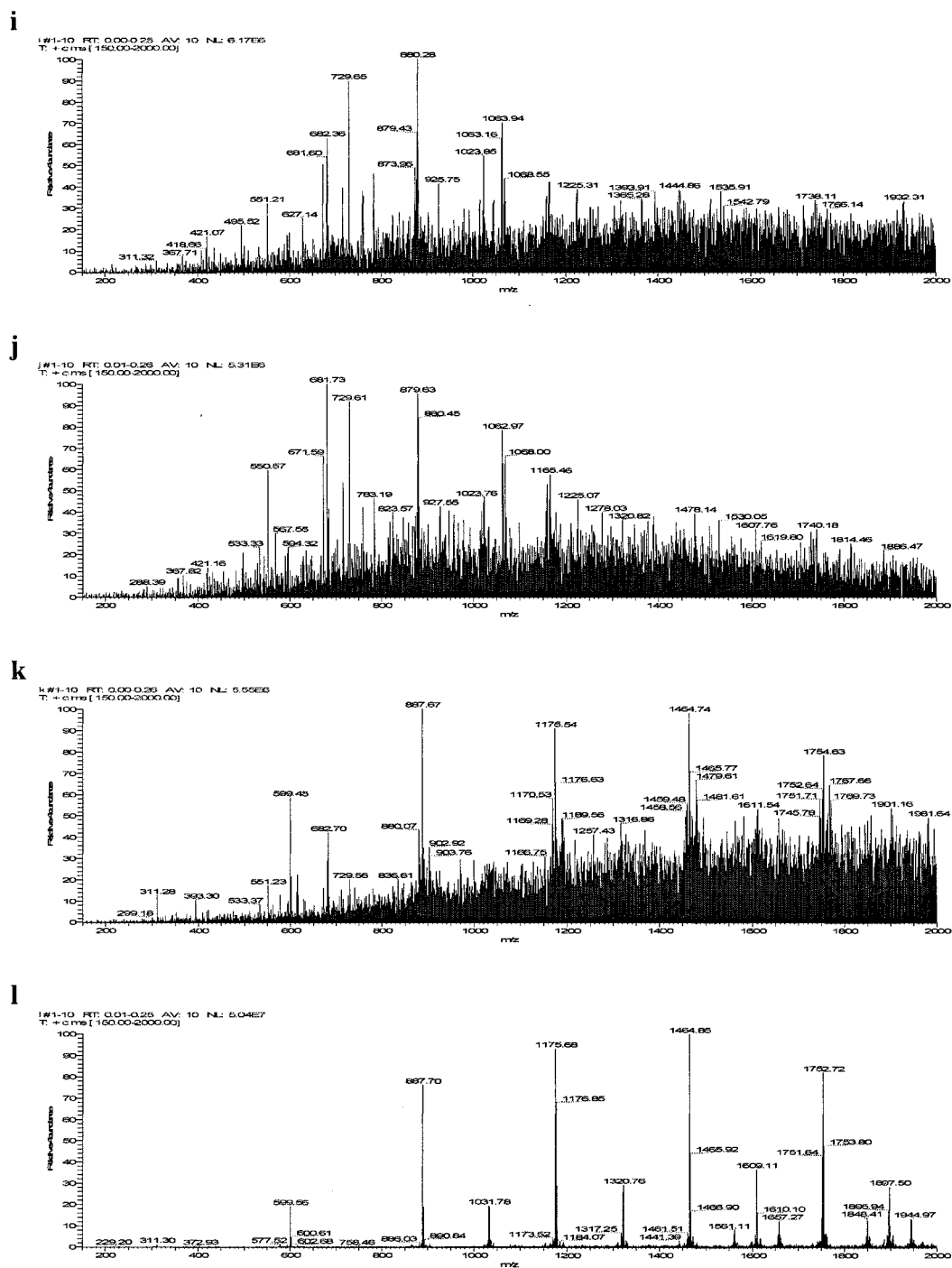
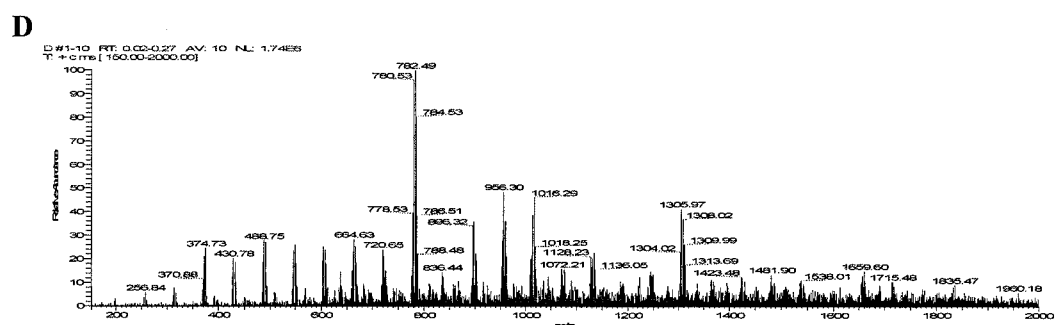
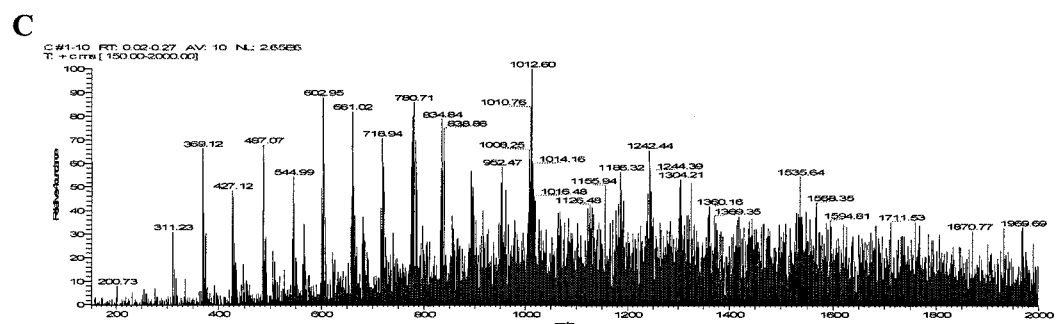
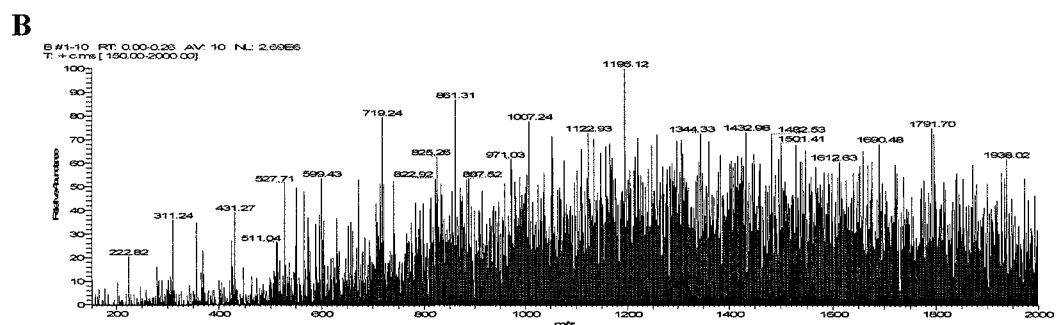
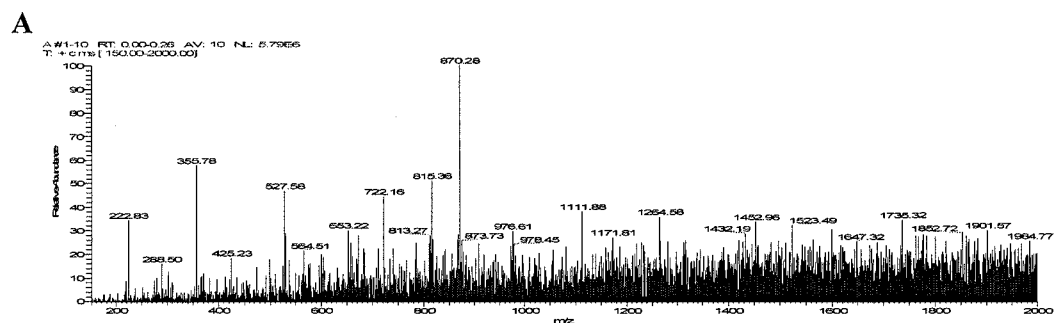
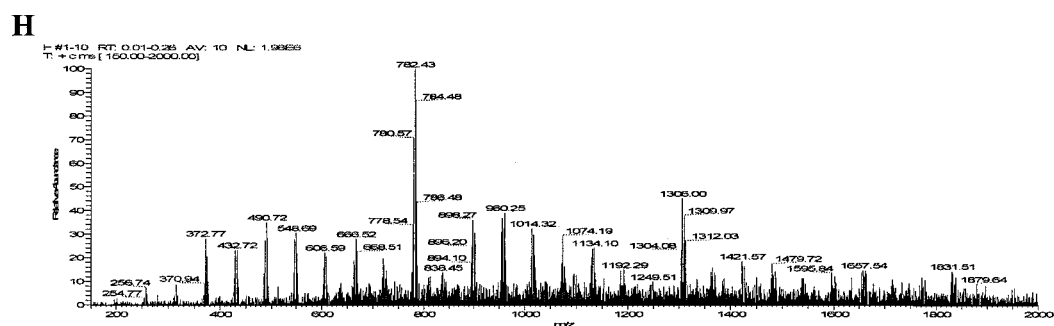
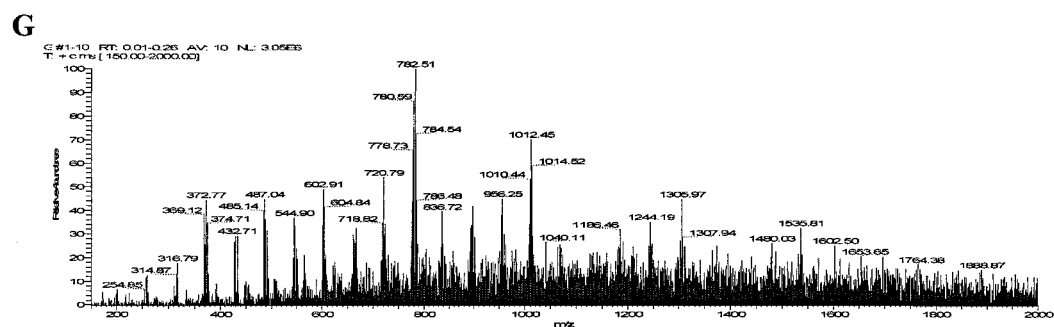
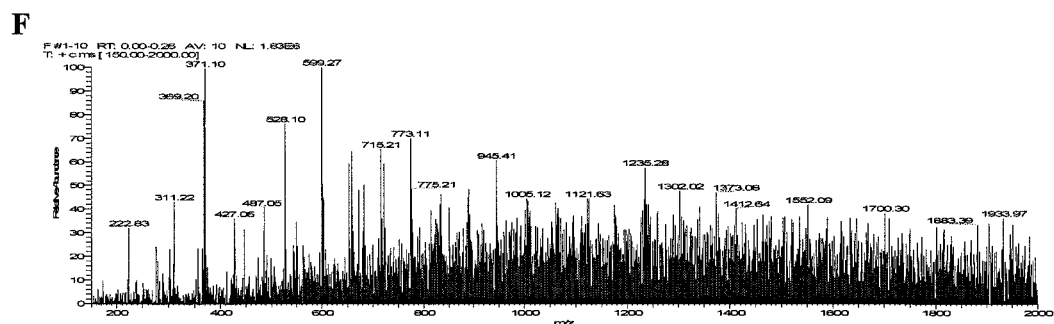
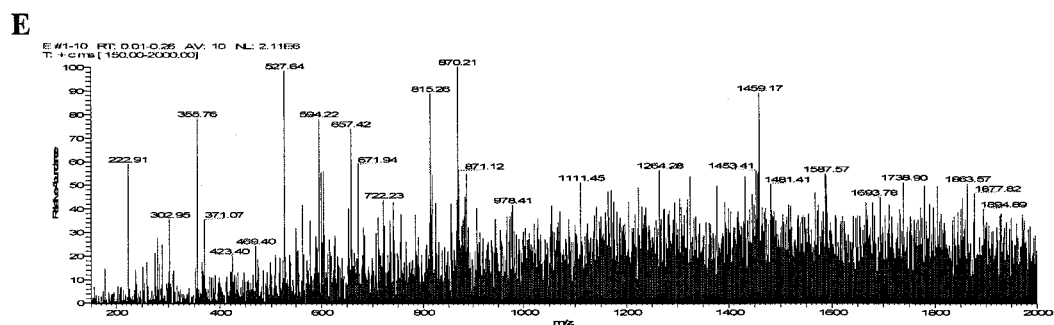


Figure F2 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using β -Galactosidase digests.

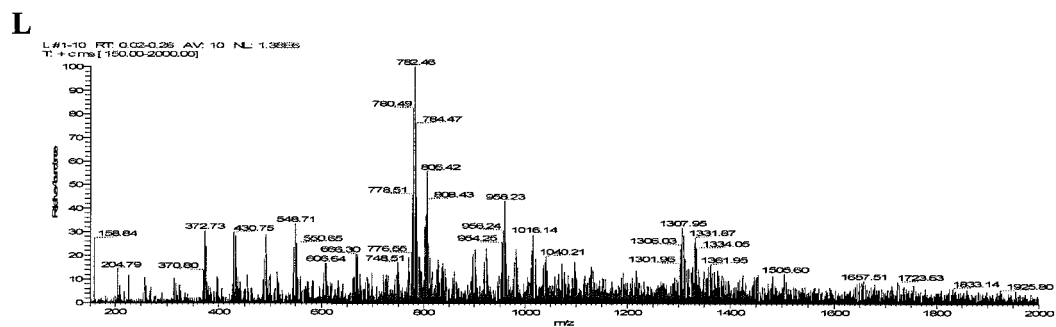
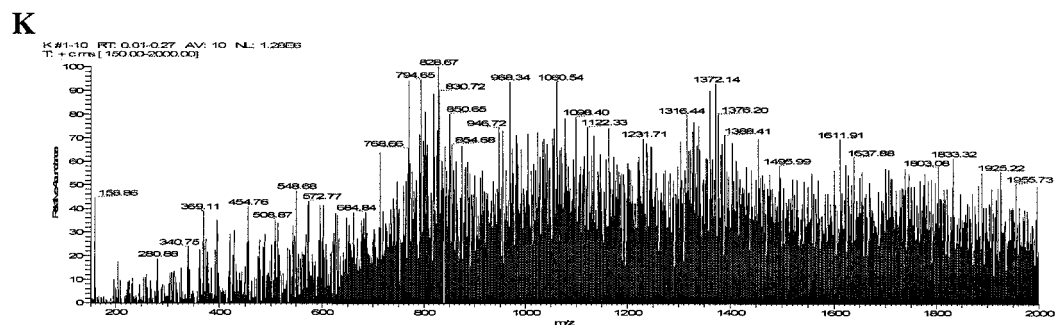
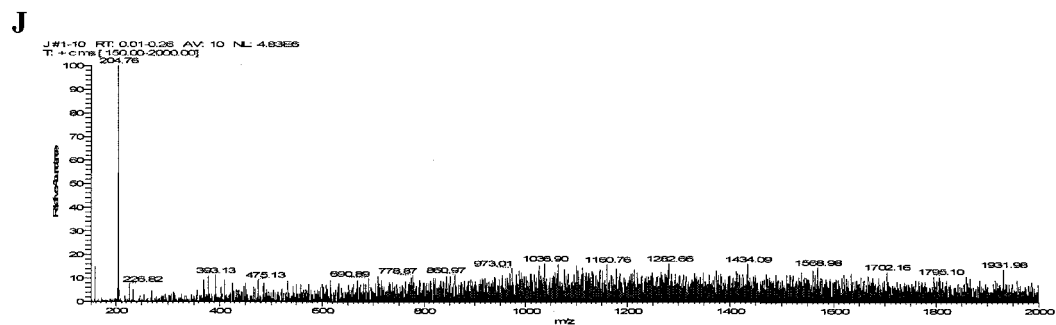
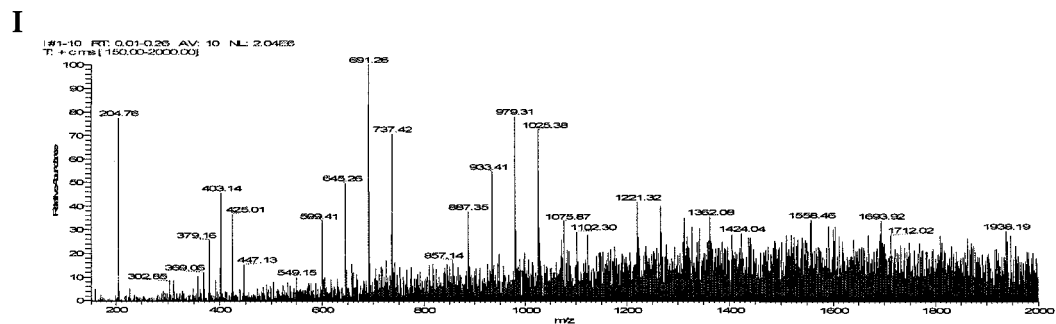
The sample conditions (A-L and a-l) are given in Table 2.10.



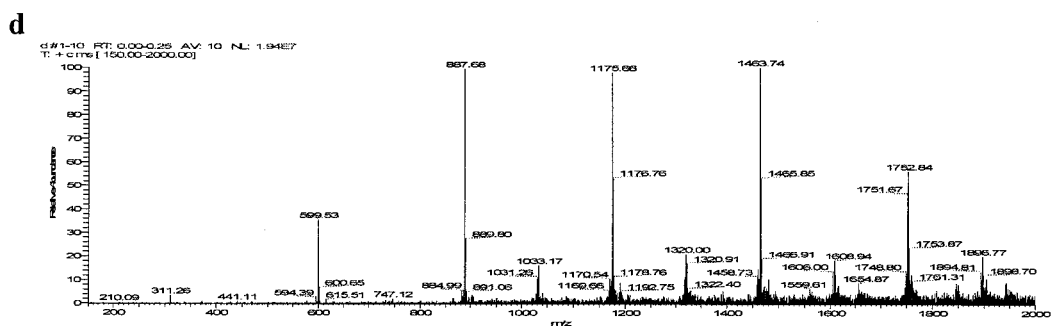
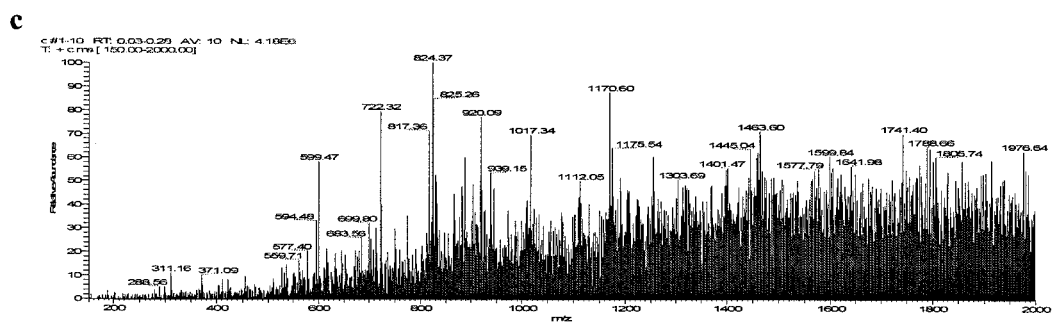
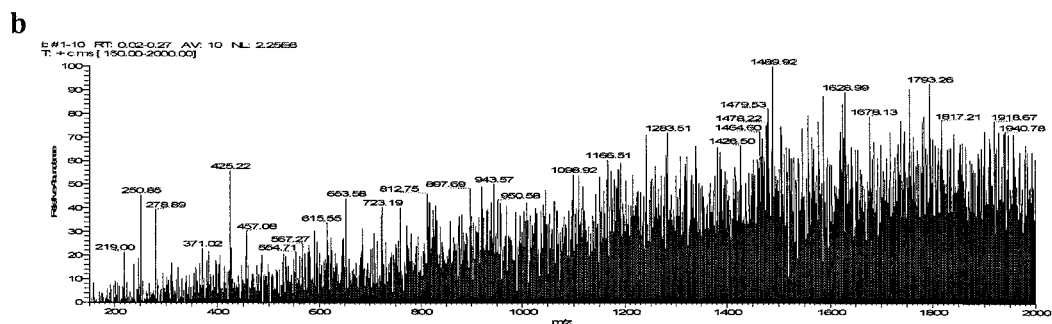
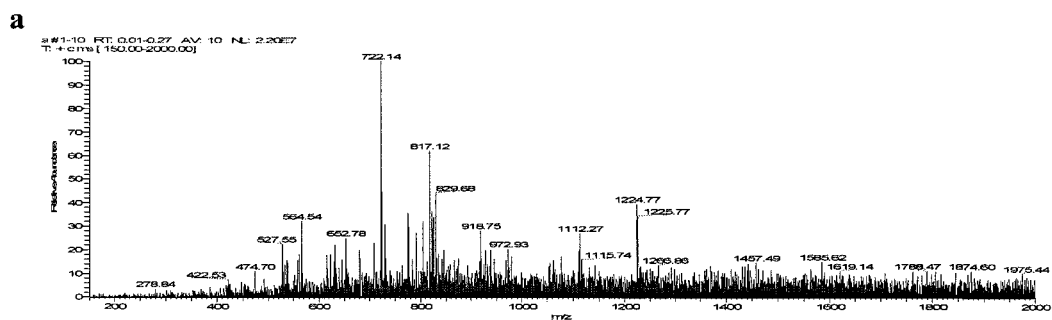
(Figure continues on next page)



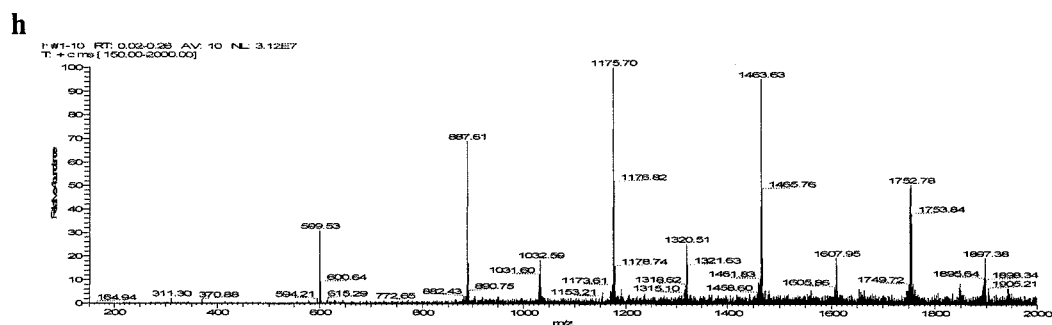
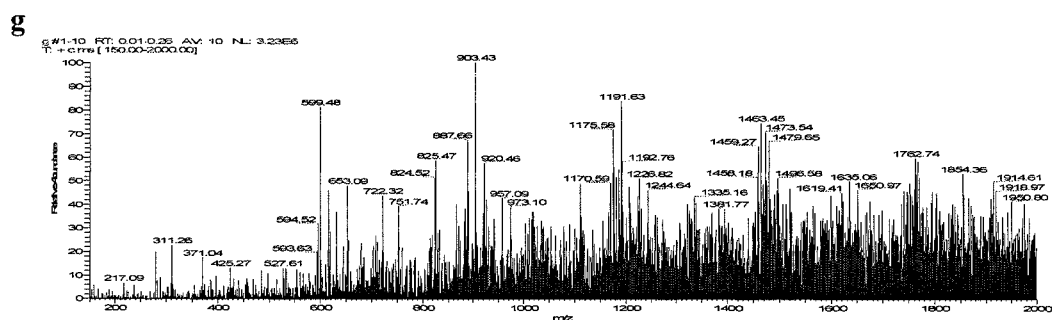
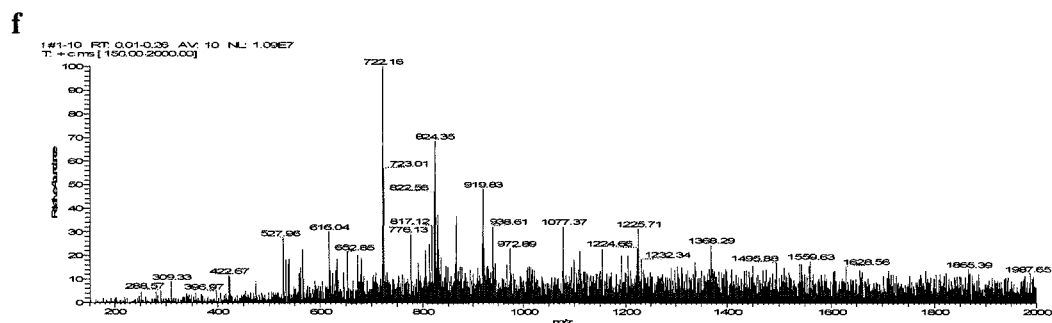
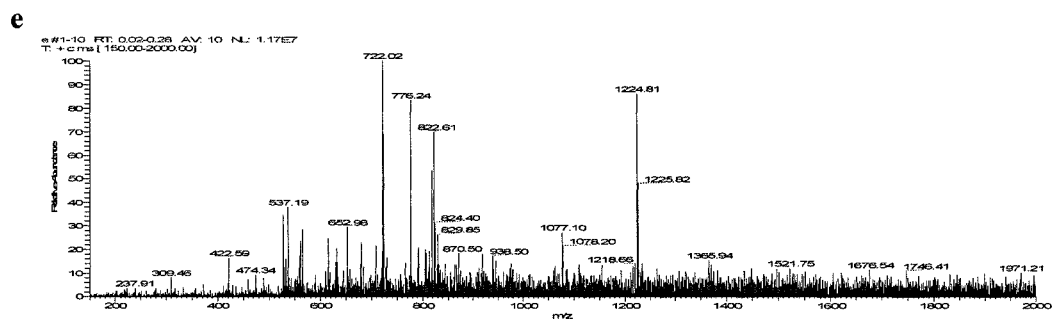
(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)

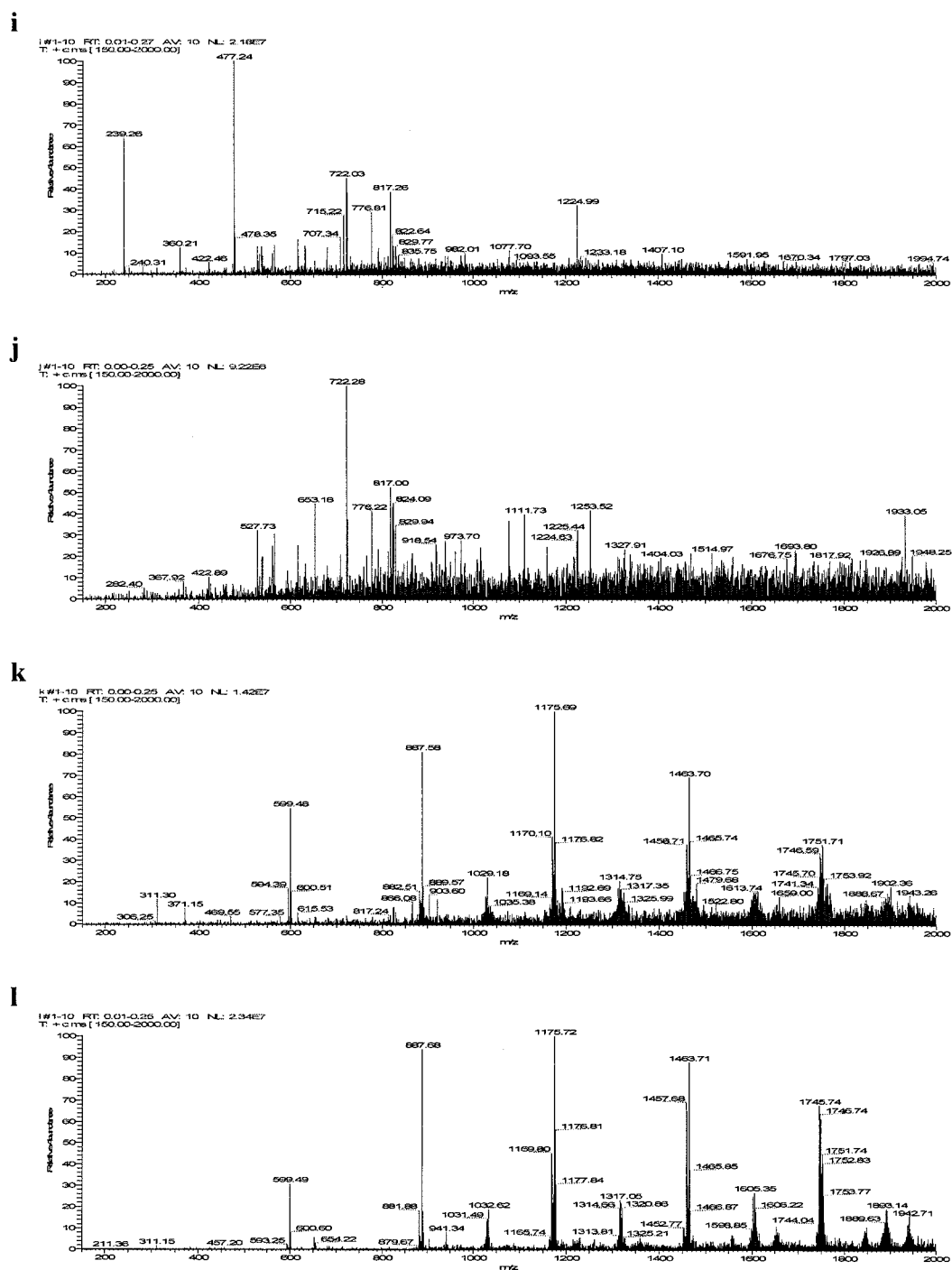
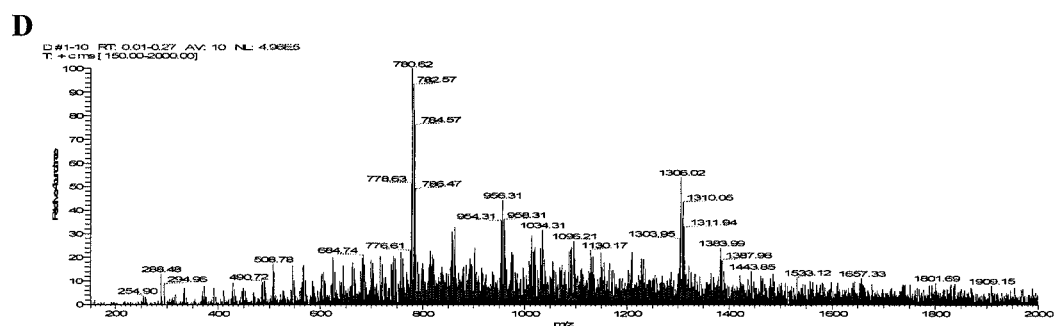
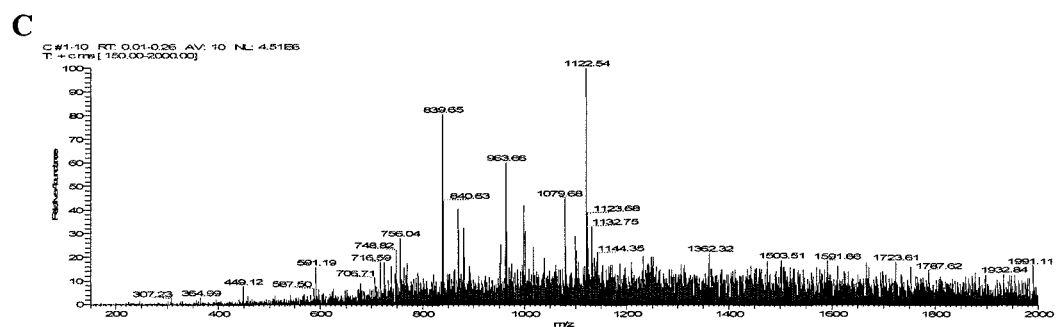
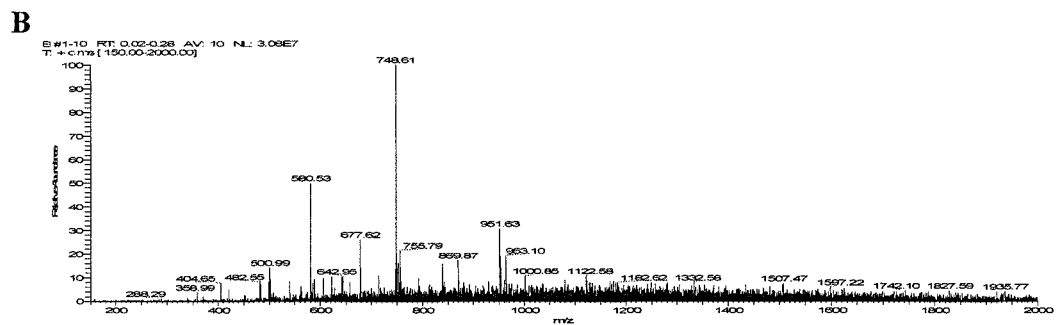
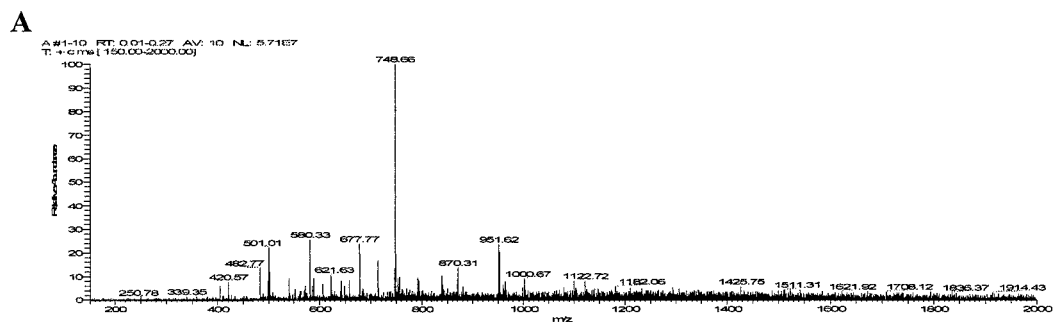
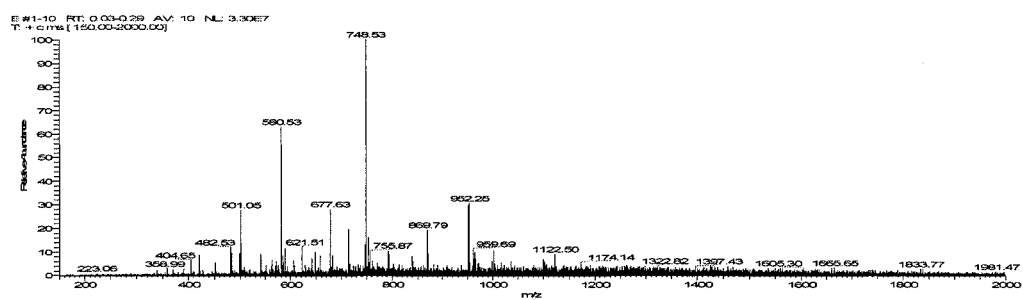
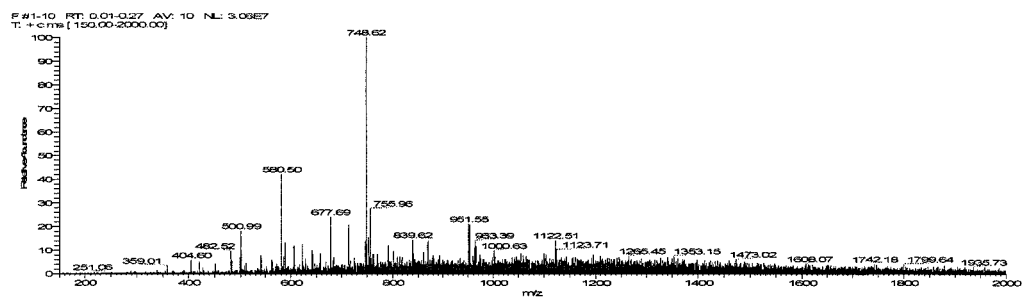
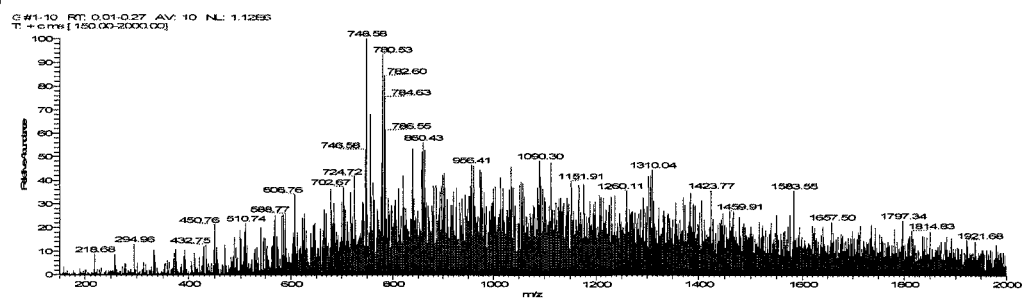
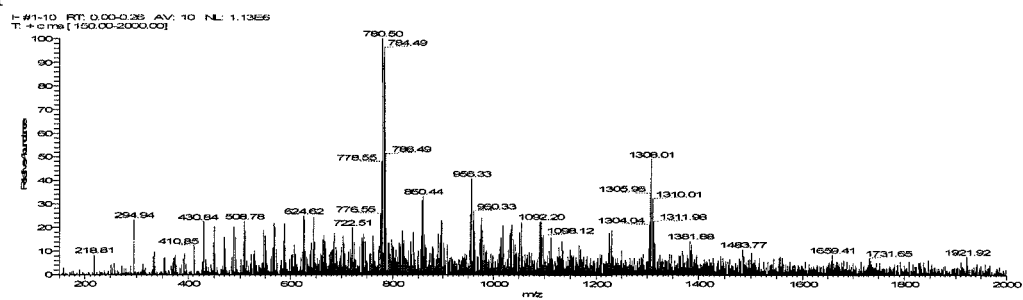


Figure F3 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using phosphorylase b digests.

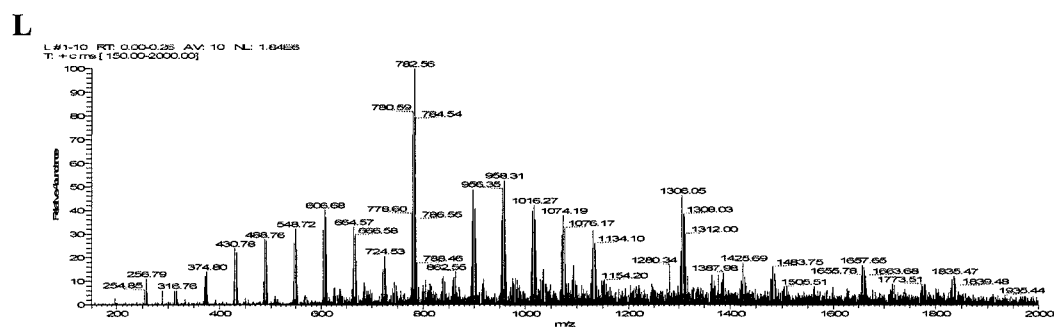
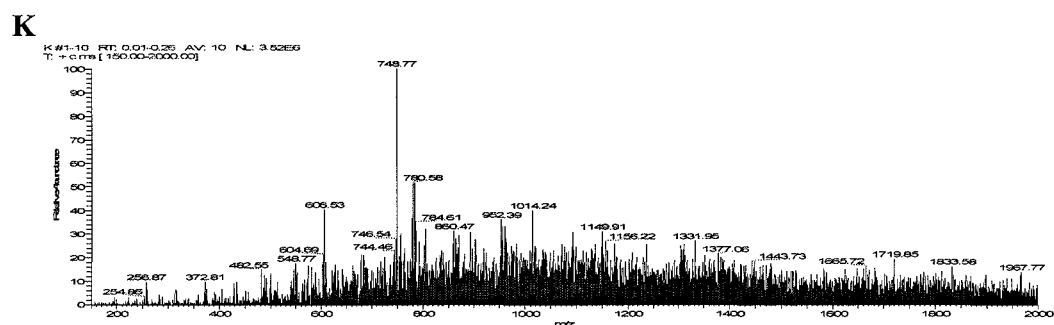
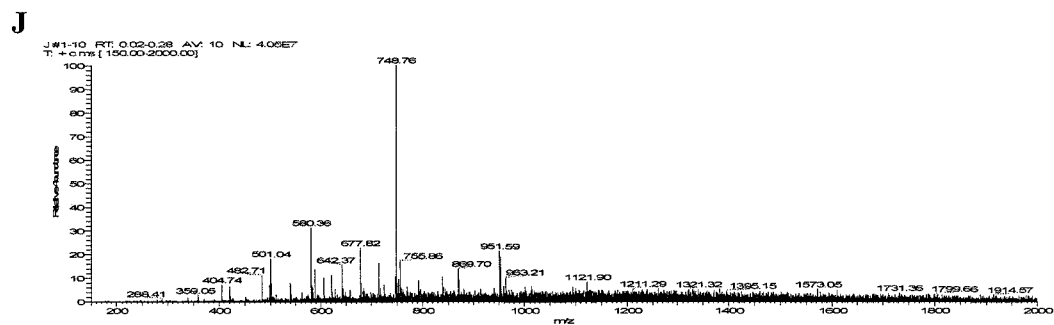
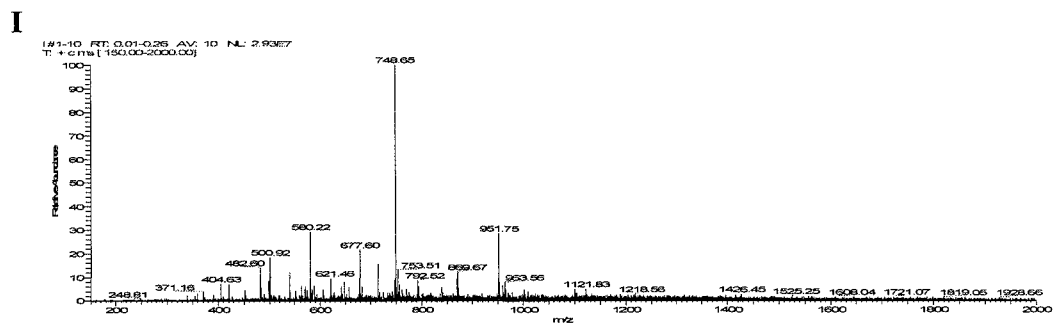
The sample conditions (A-L and a-l) are given in Table 2.10.



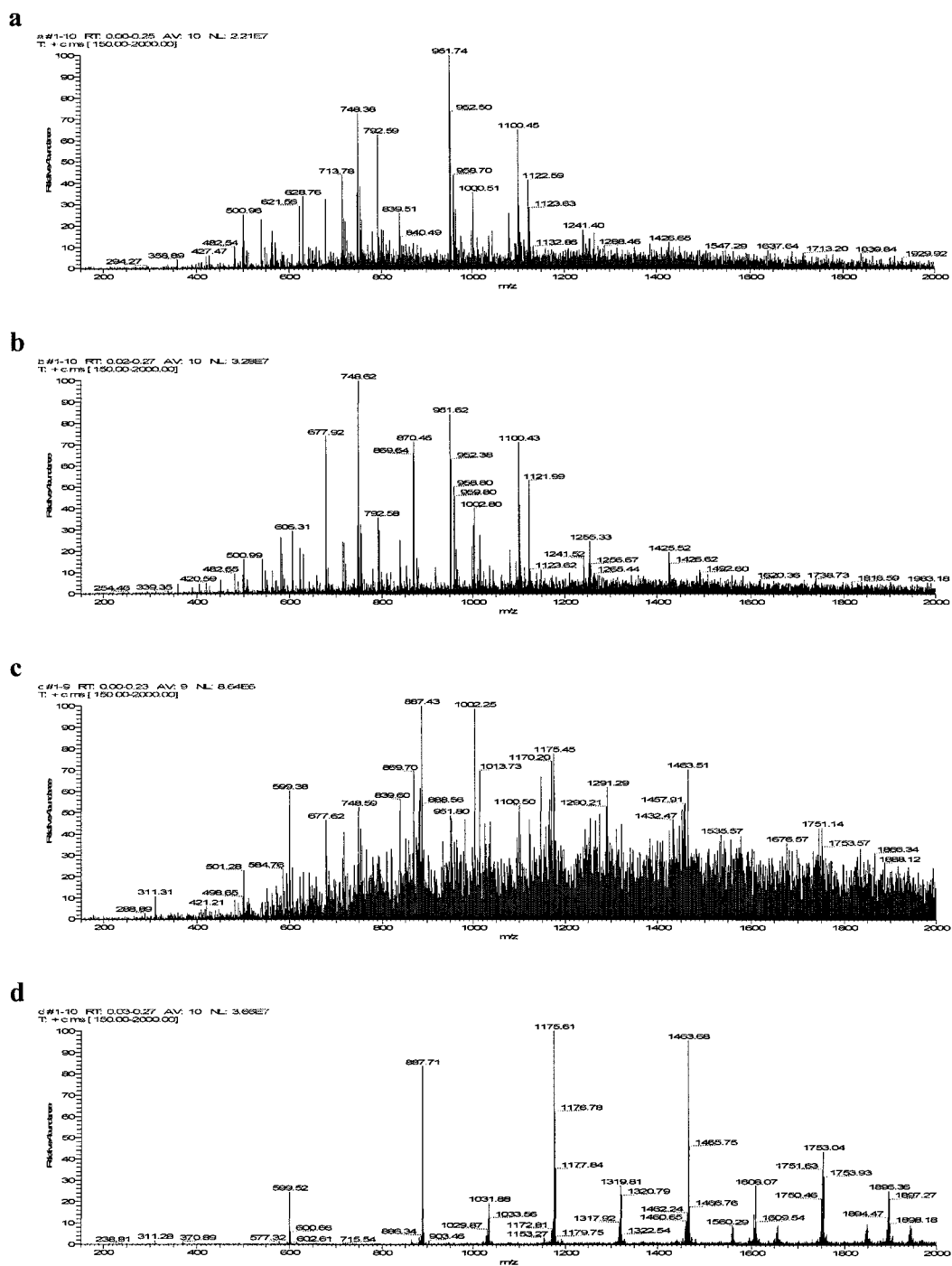
(Figure continues on next page)

E**F****G****H**

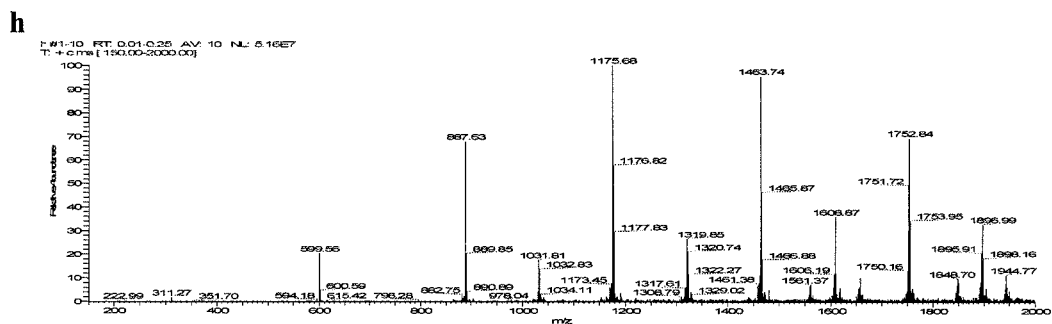
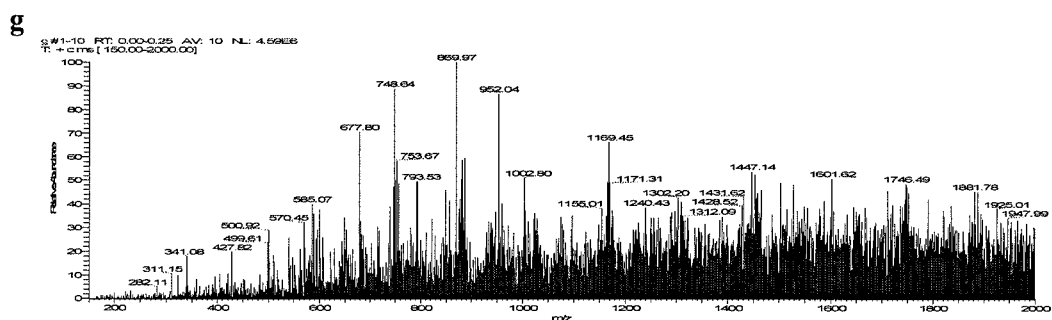
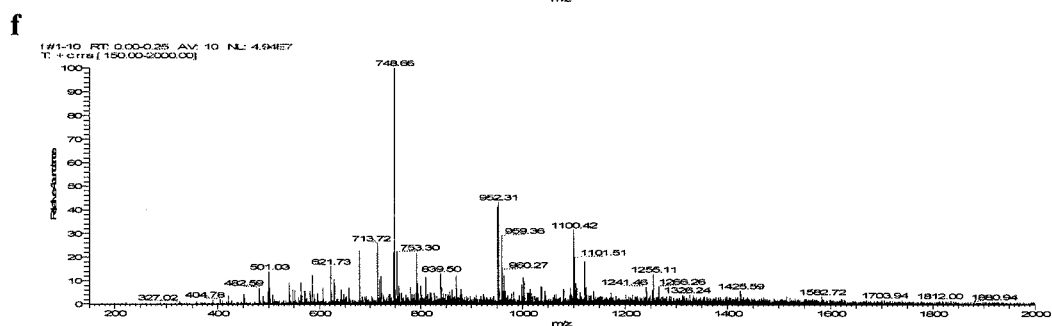
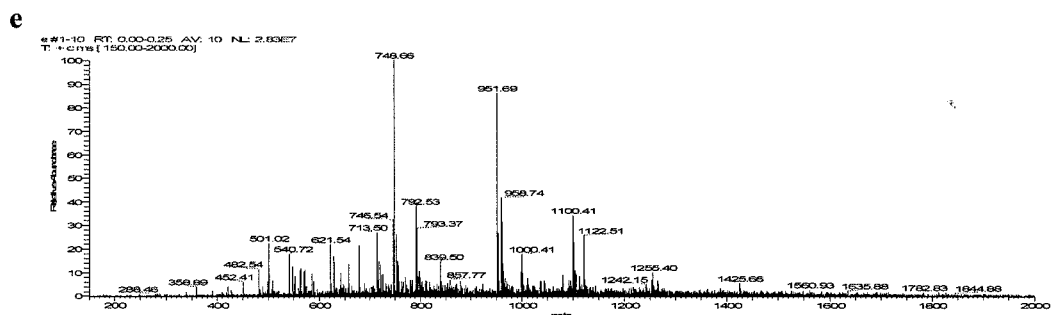
(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)

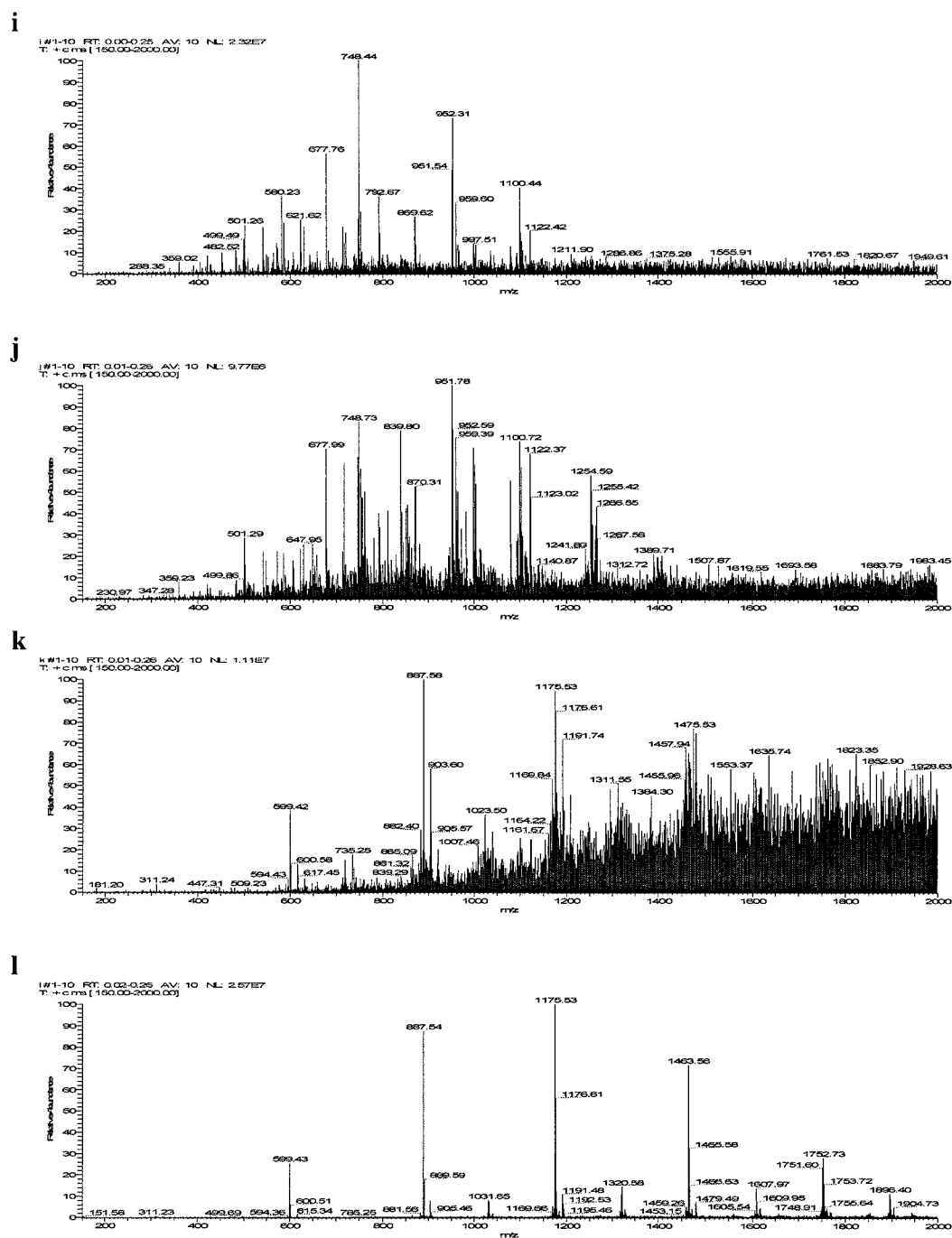
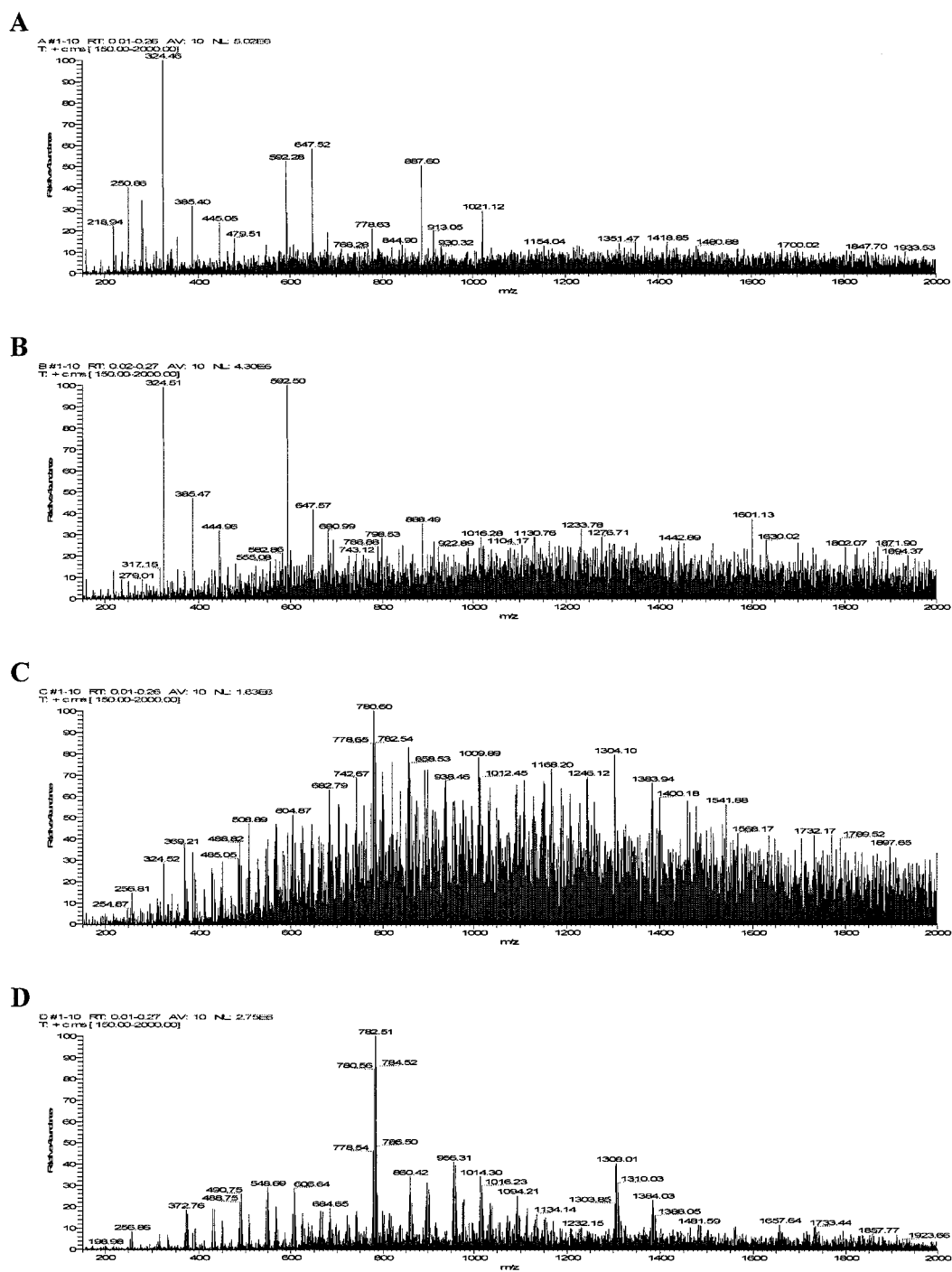
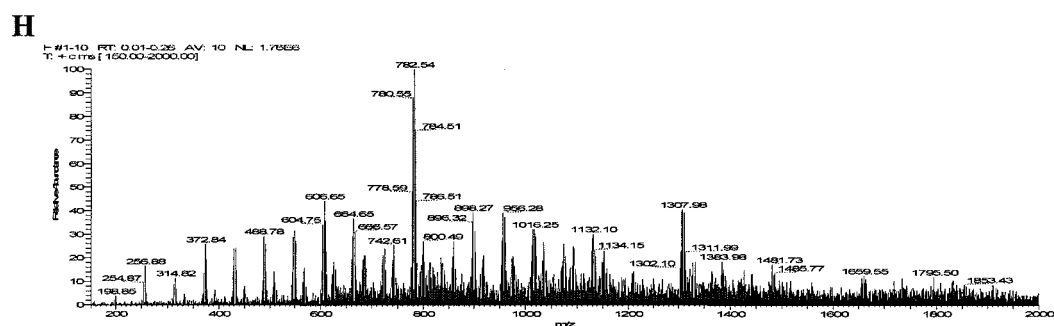
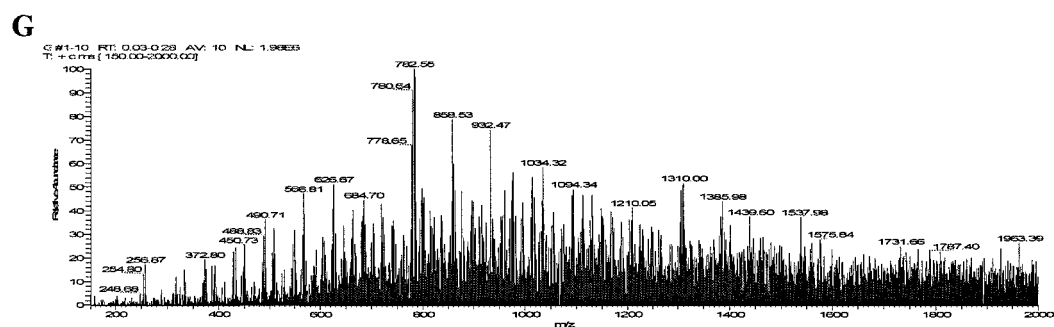
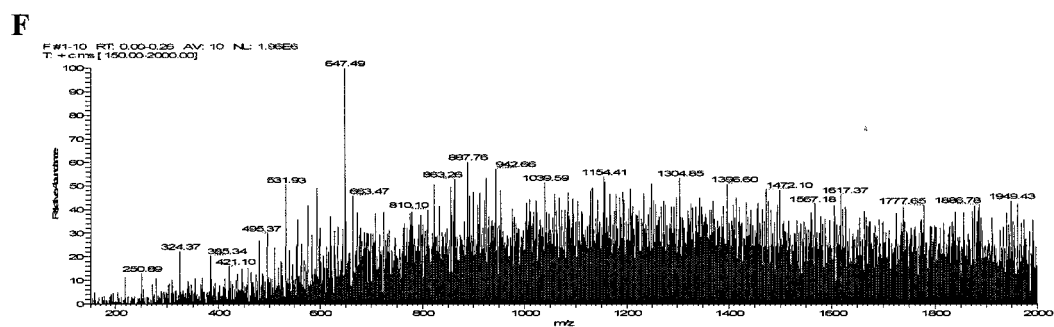
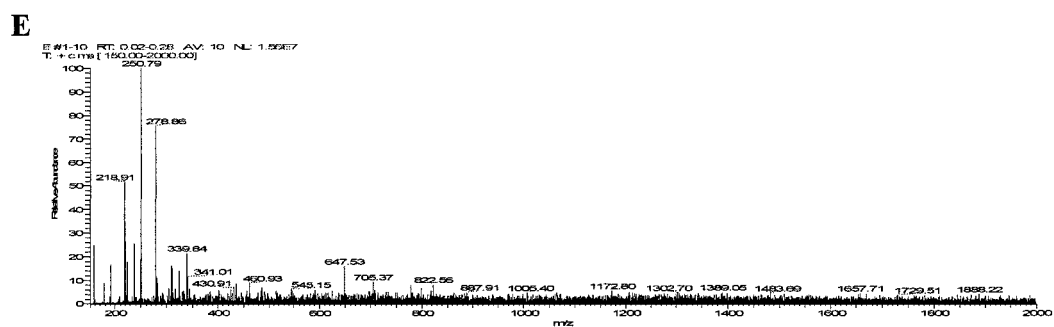


Figure F4 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using L-glutamic dehydrogenase digests.

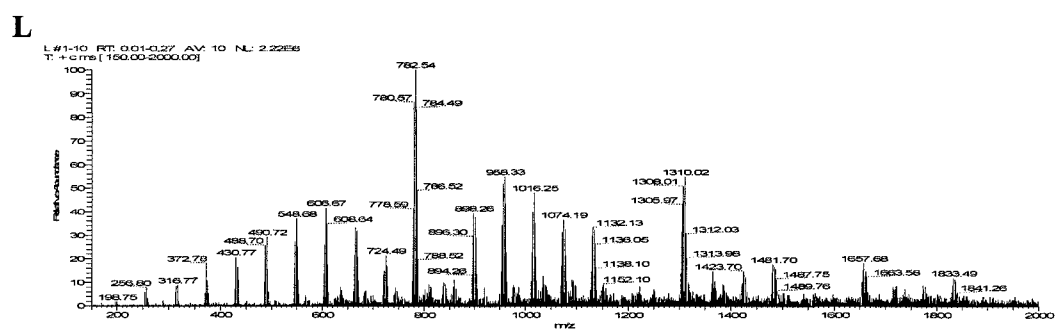
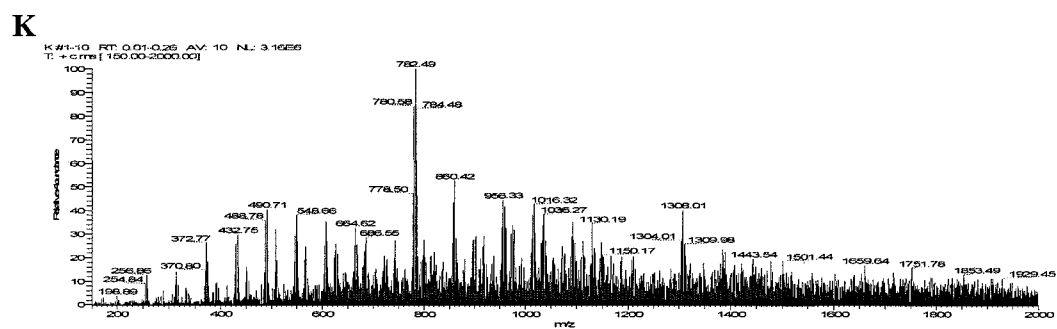
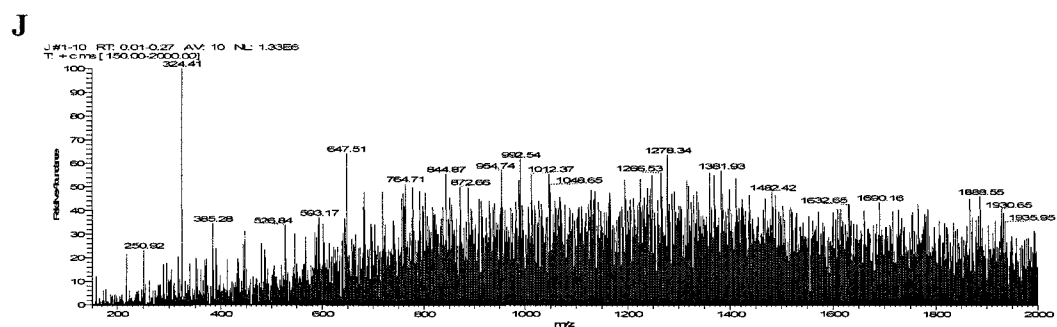
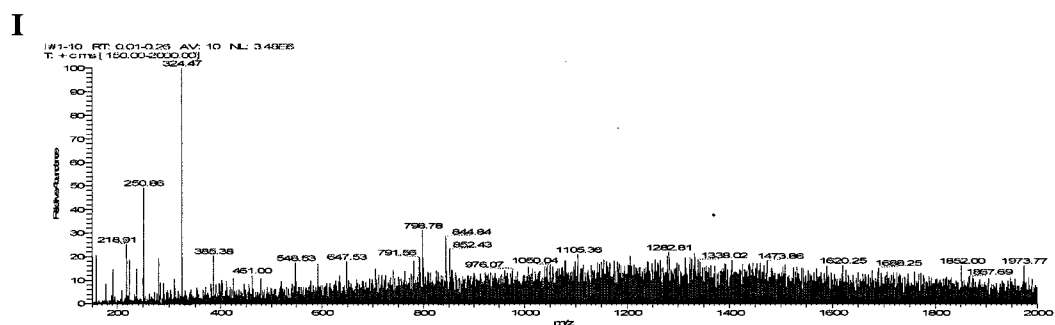
The sample conditions (A-L and a-l) are given in Table 2.10.



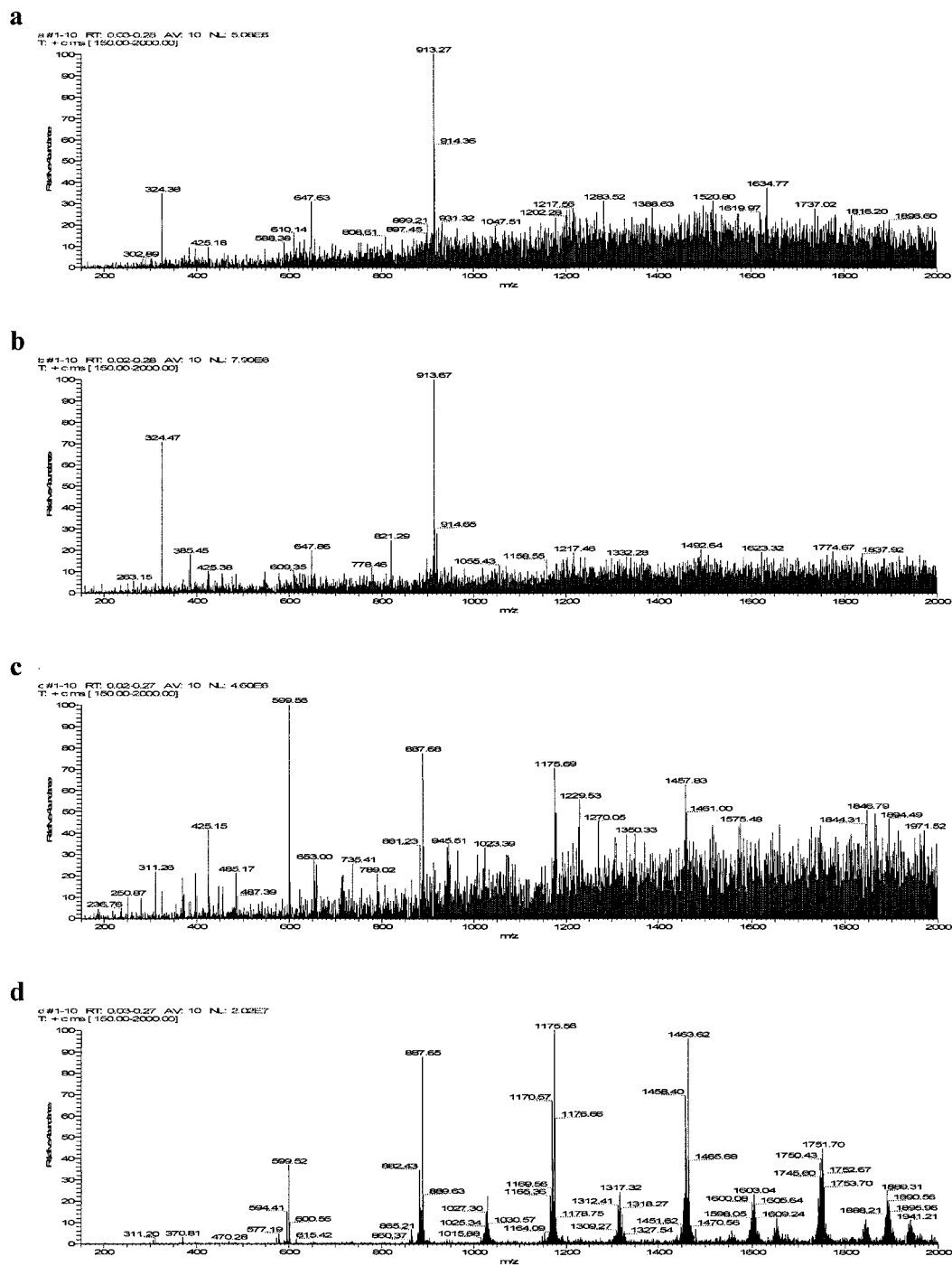
(Figure continues on next page)

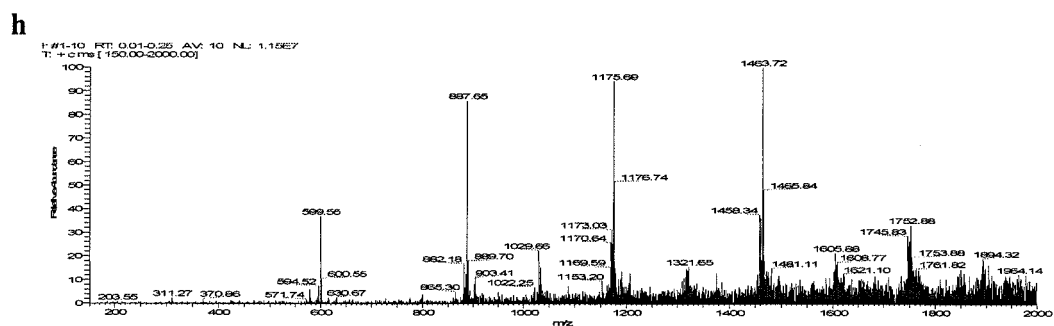
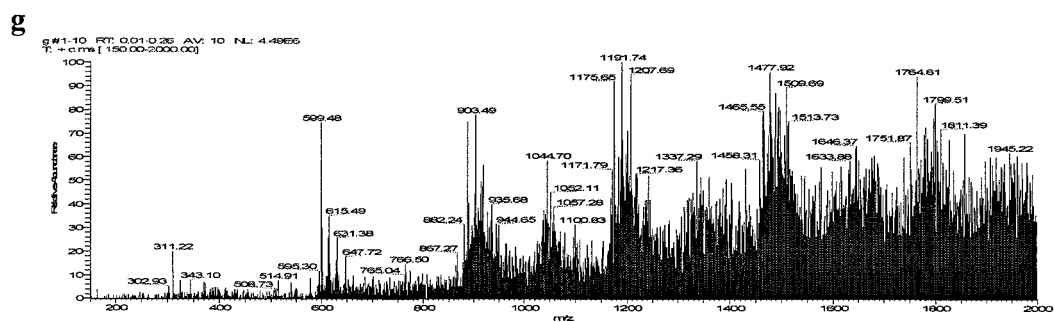
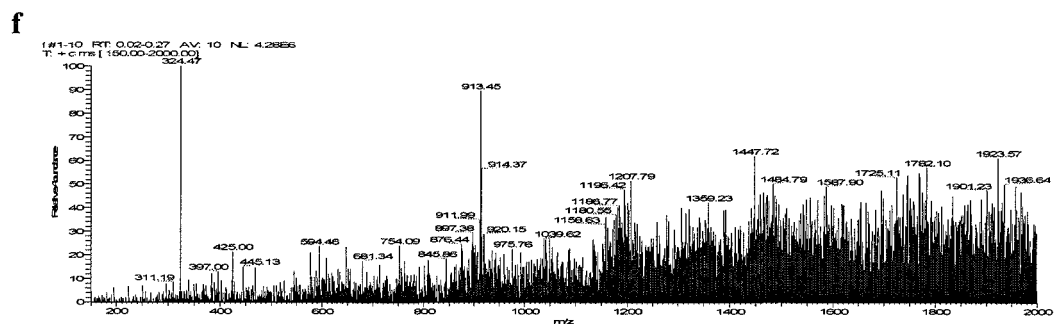
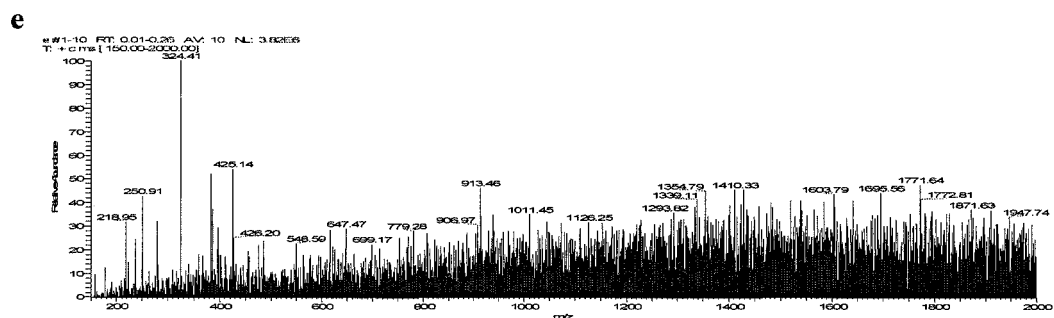


(Figure continues on next page)



(Figure continues on next page)





(Figure continues on next page)

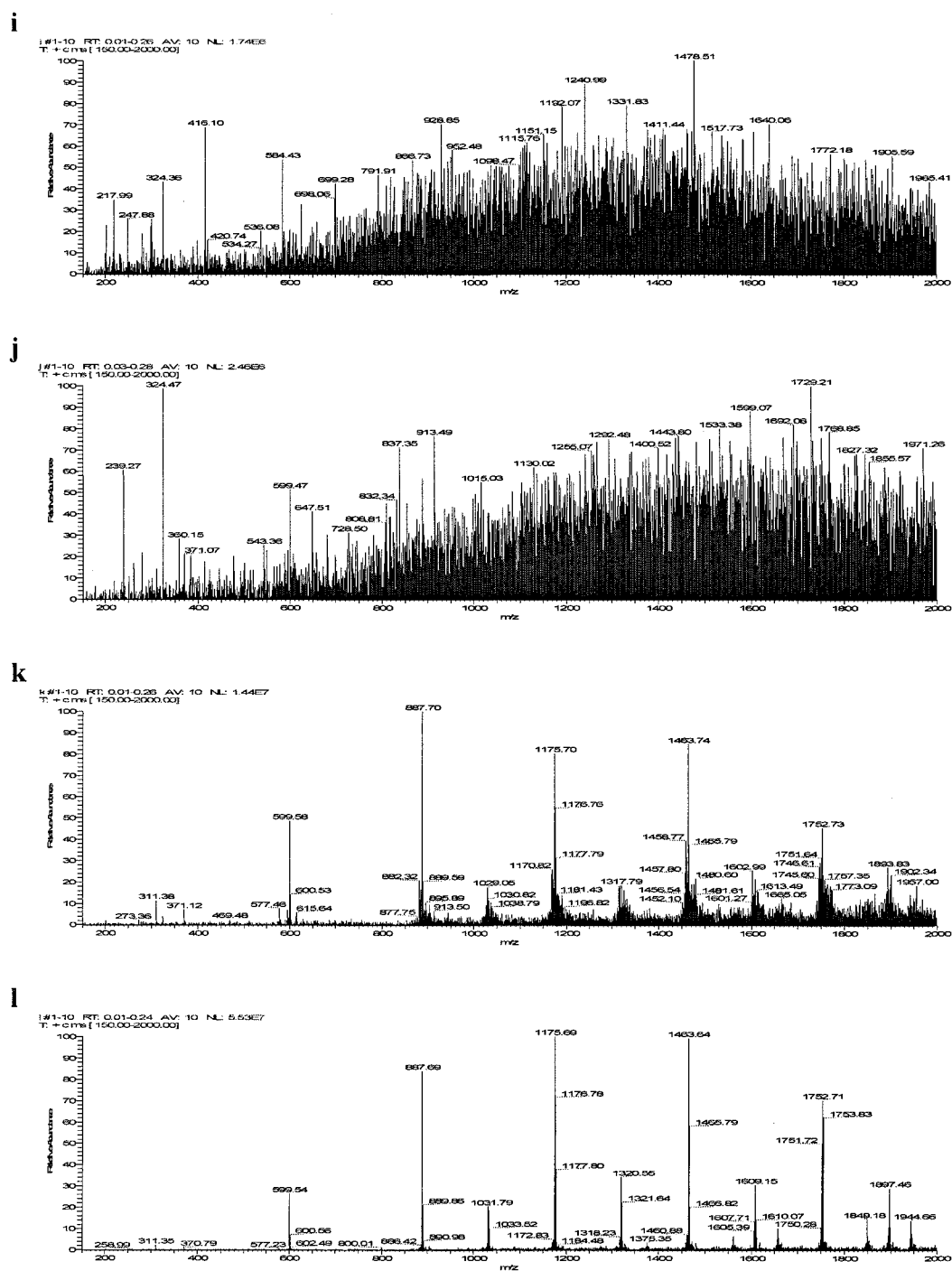
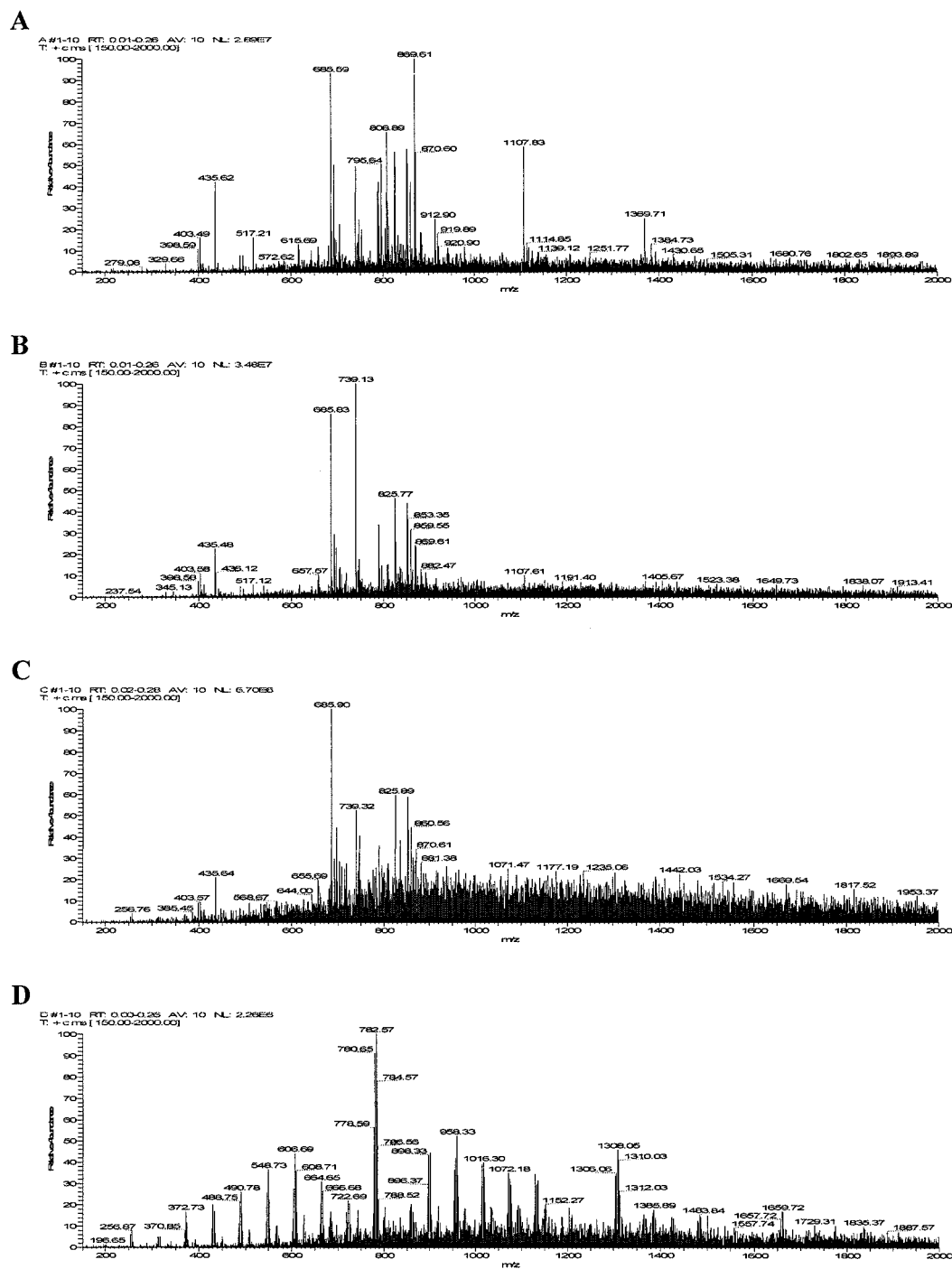
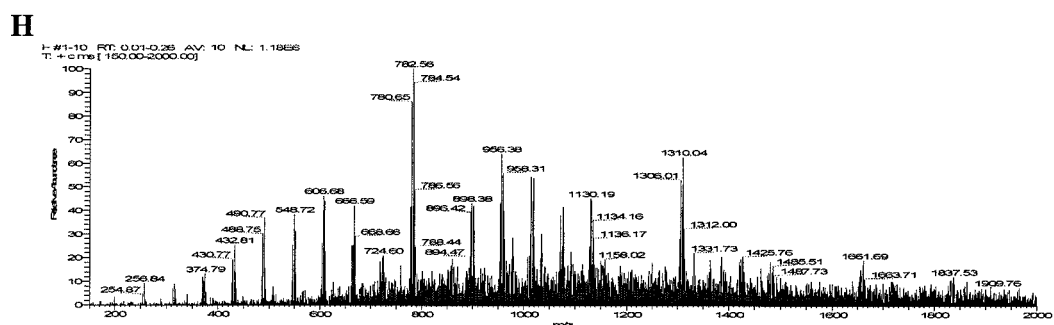
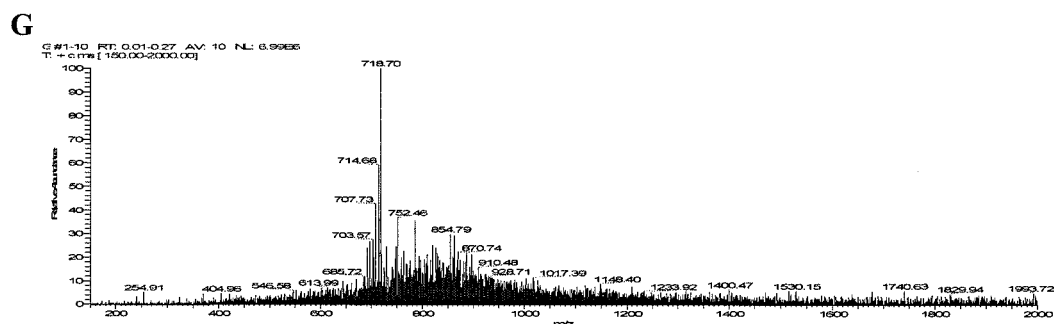
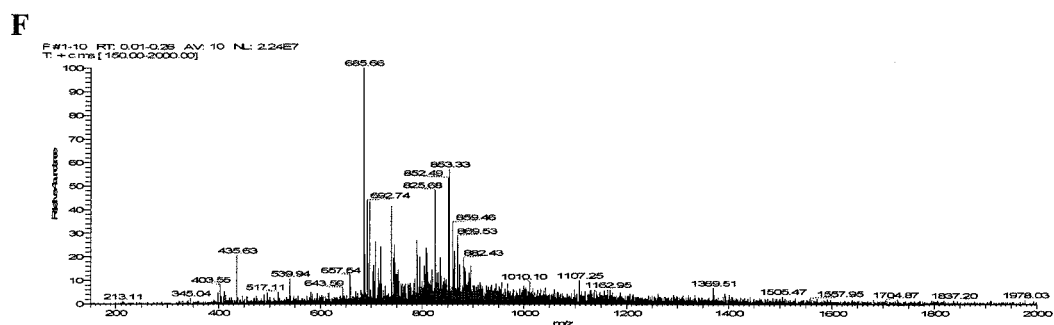
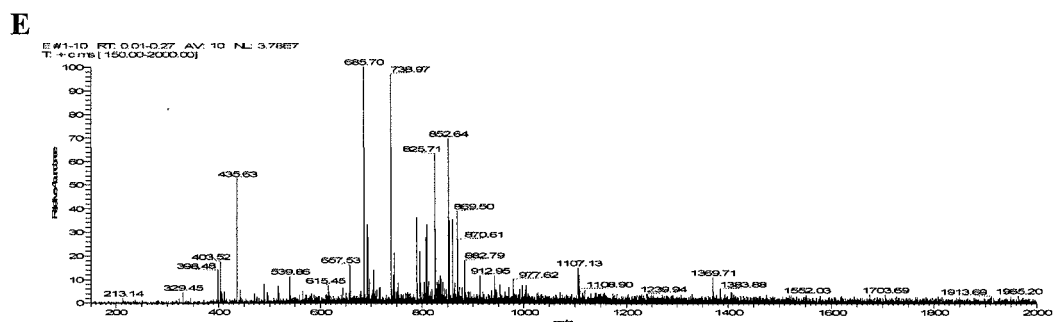


Figure F5 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using ovalbumin digests.

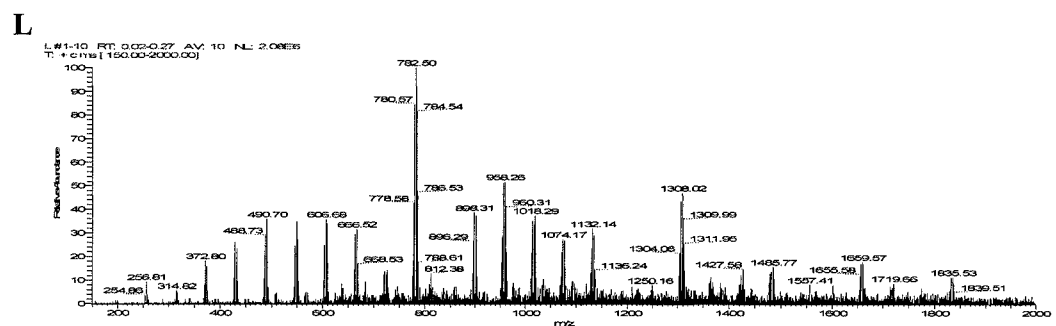
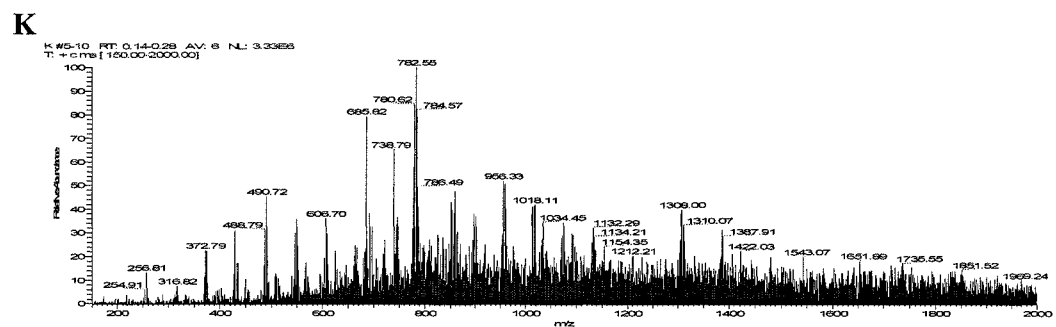
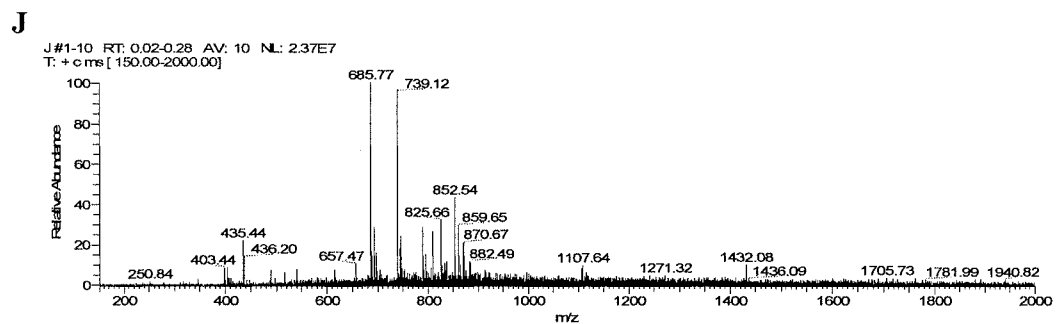
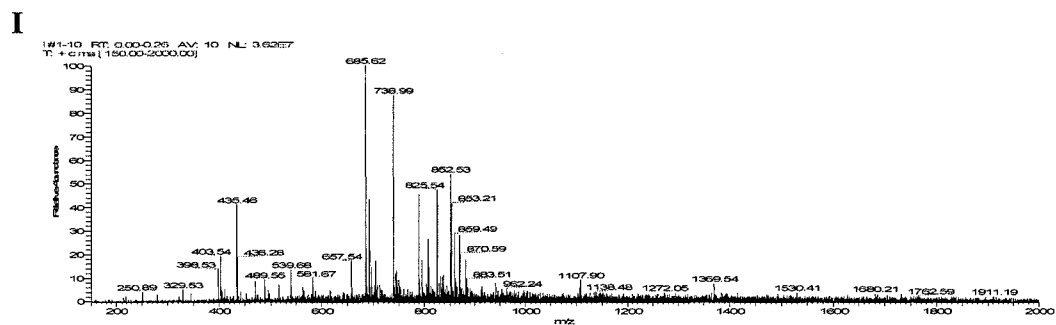
The sample conditions (A-L and a-l) are given in Table 2.10.



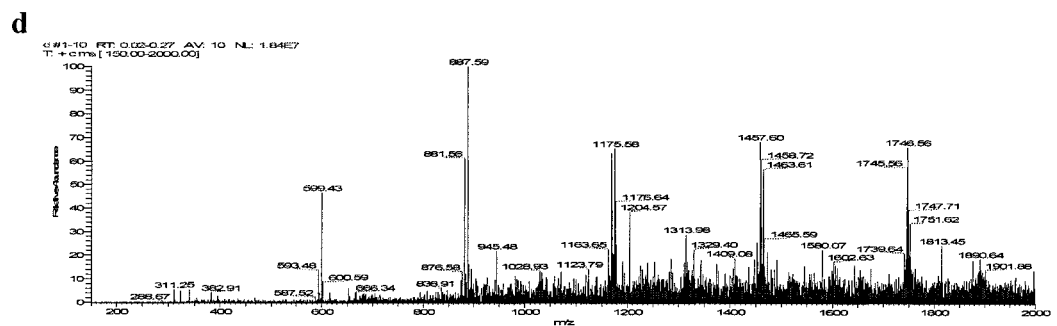
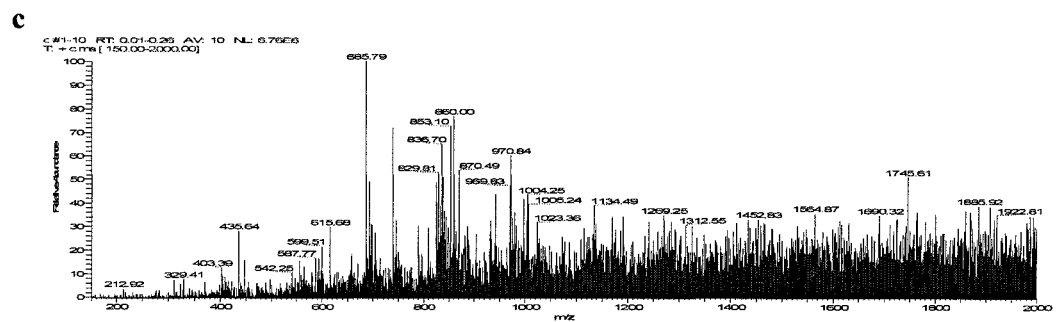
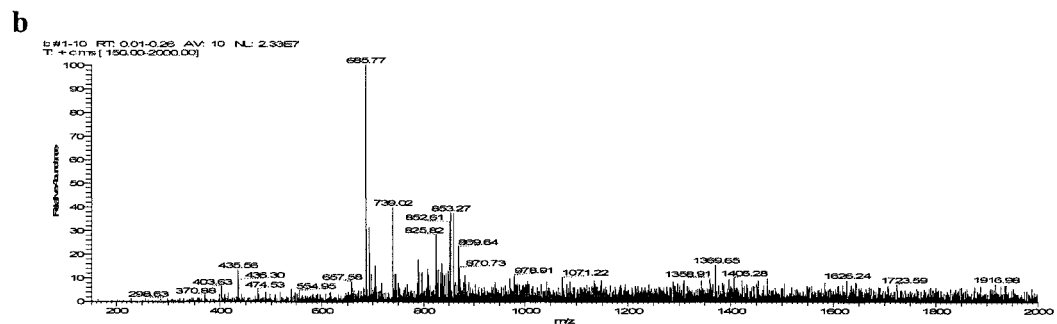
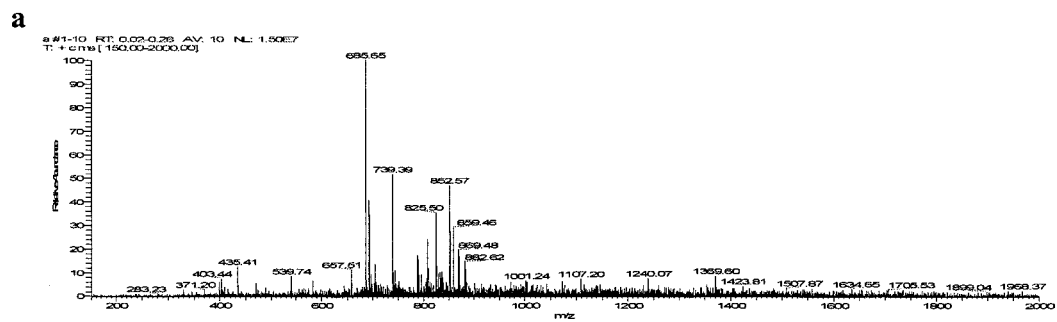
(Figure continues on next page)



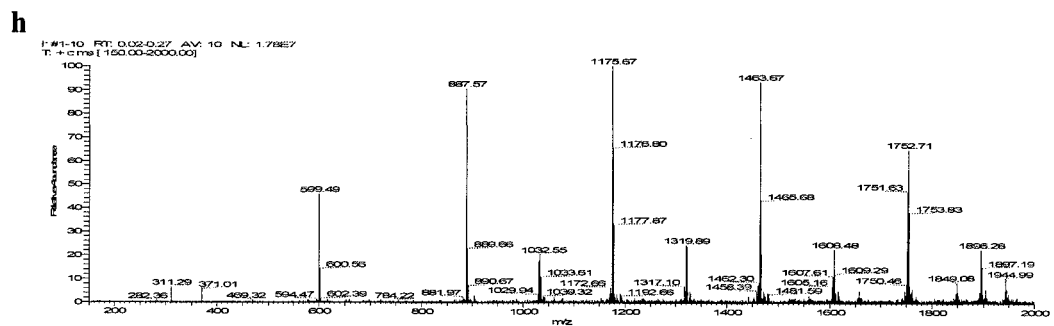
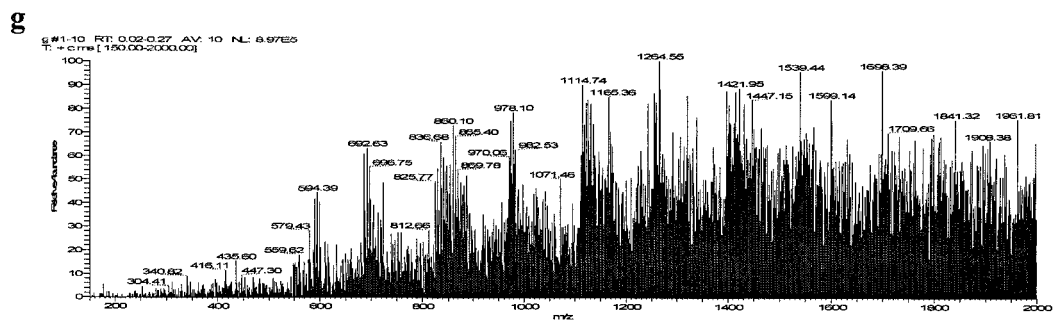
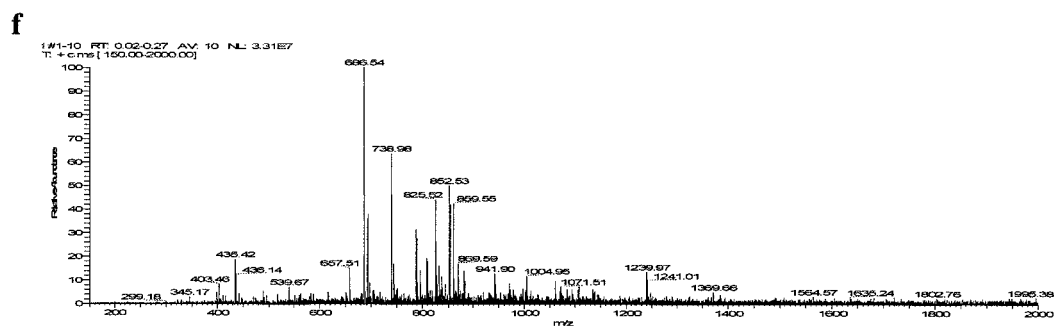
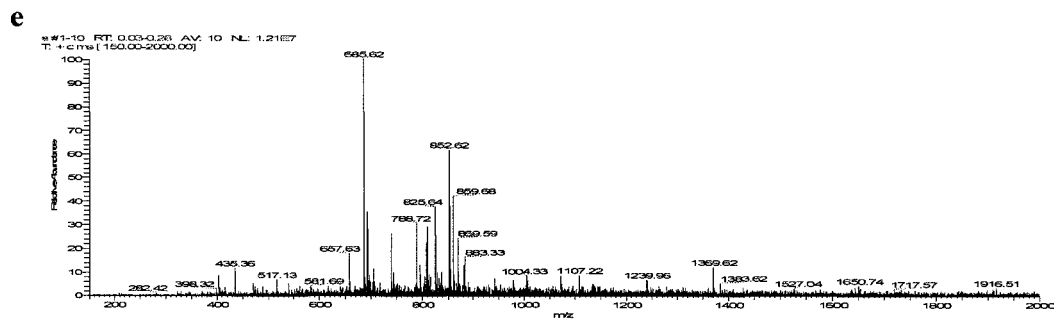
(Figure continues on next page)



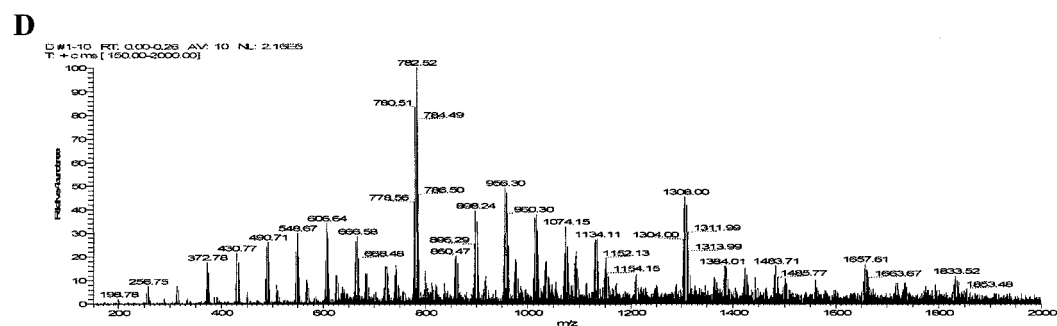
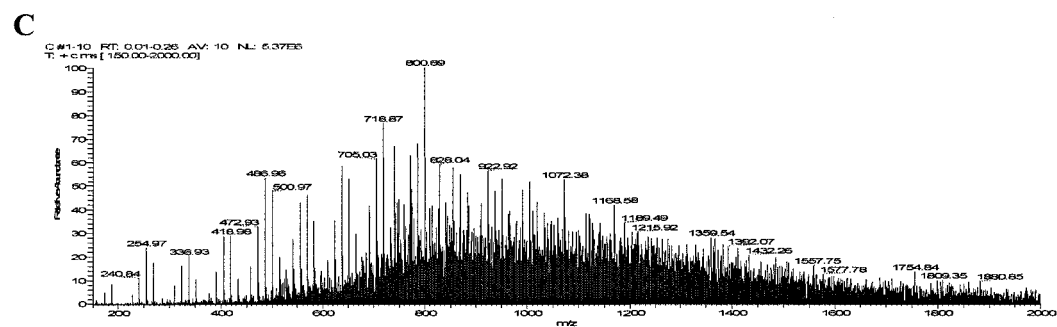
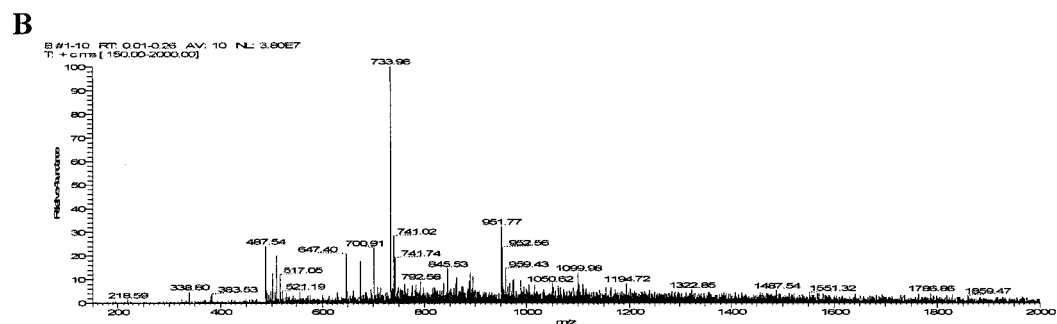
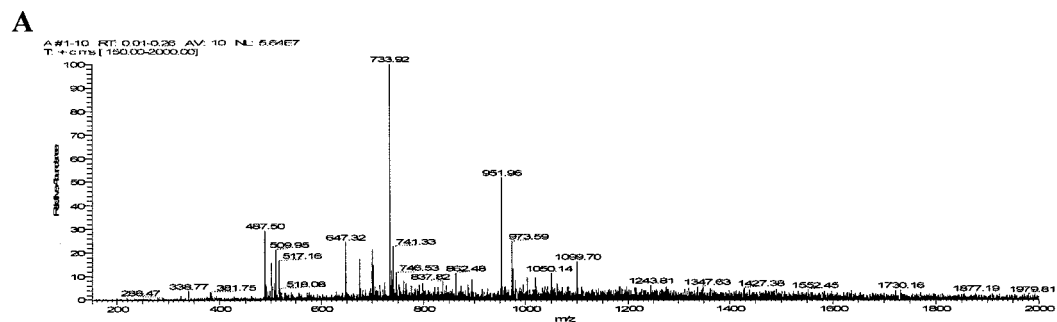
(Figure continues on next page)



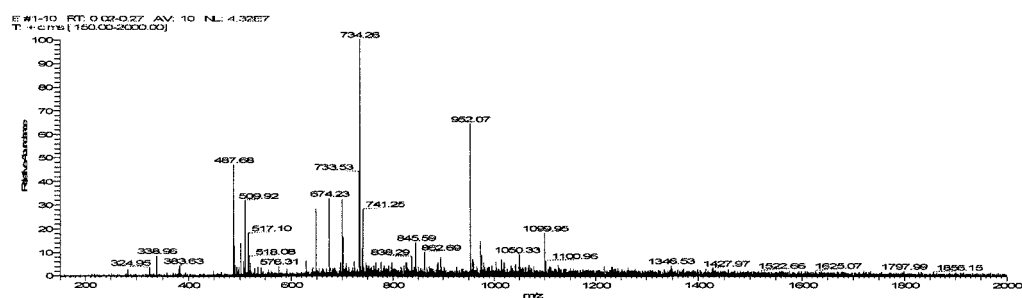
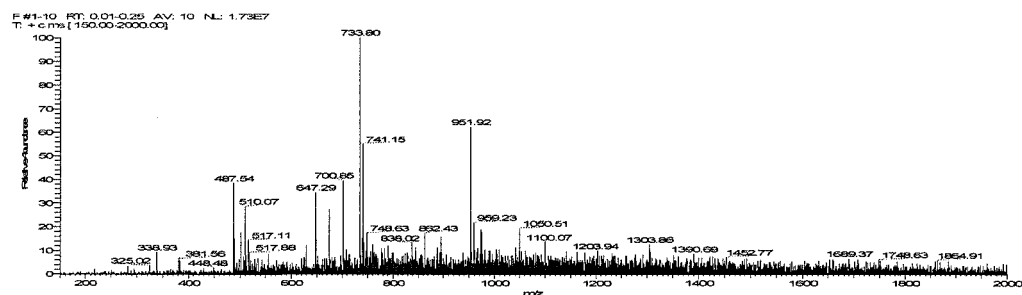
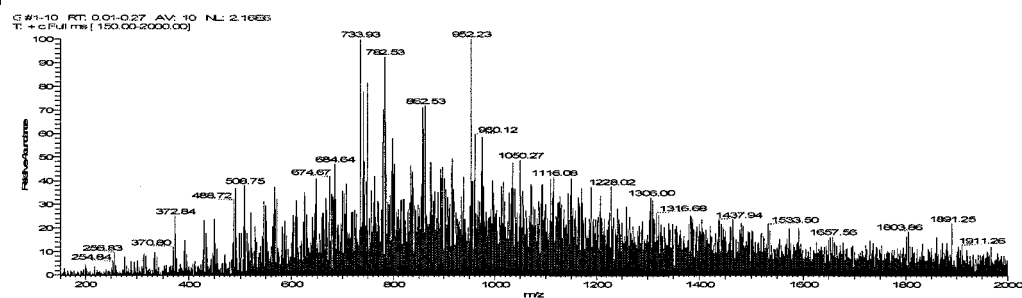
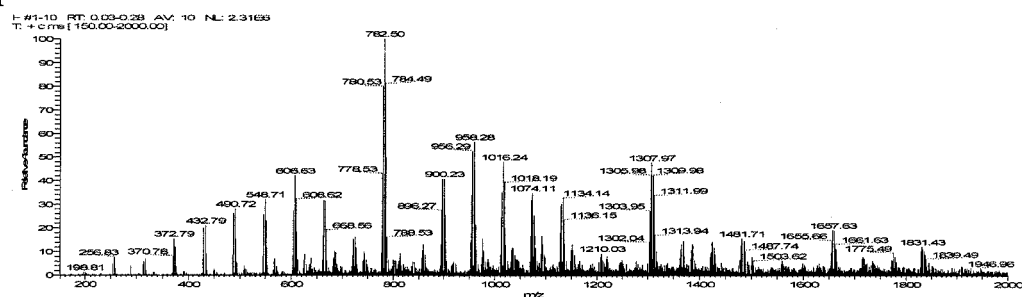
(Figure continues on next page)



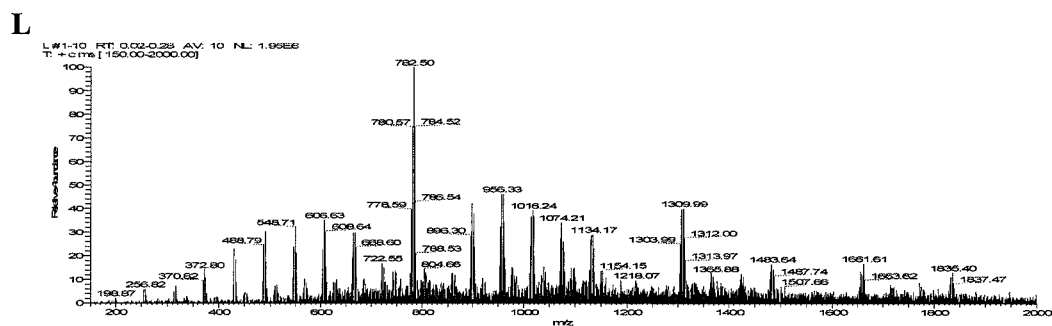
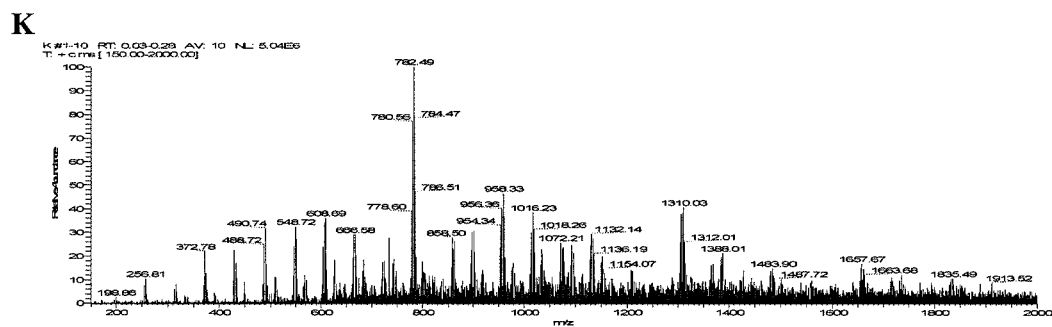
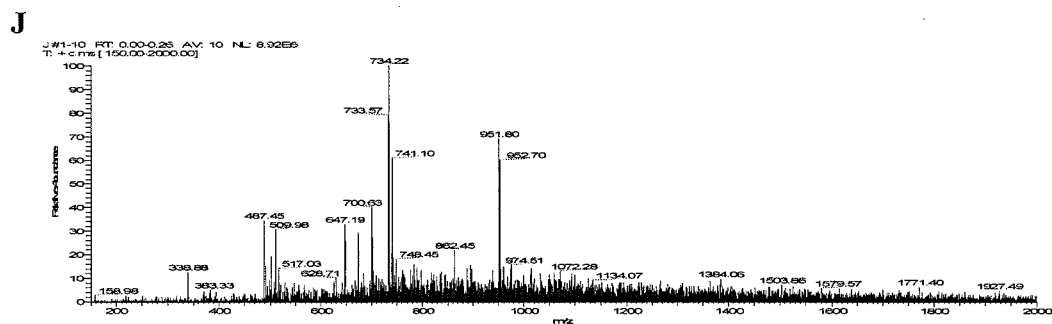
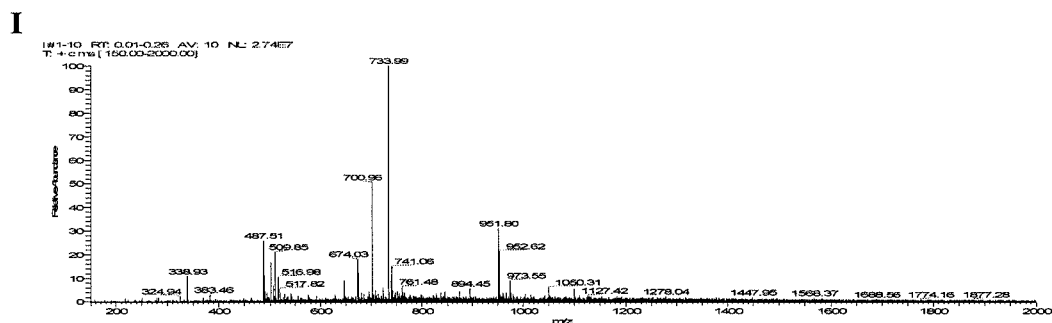
(Figure continues on next page)



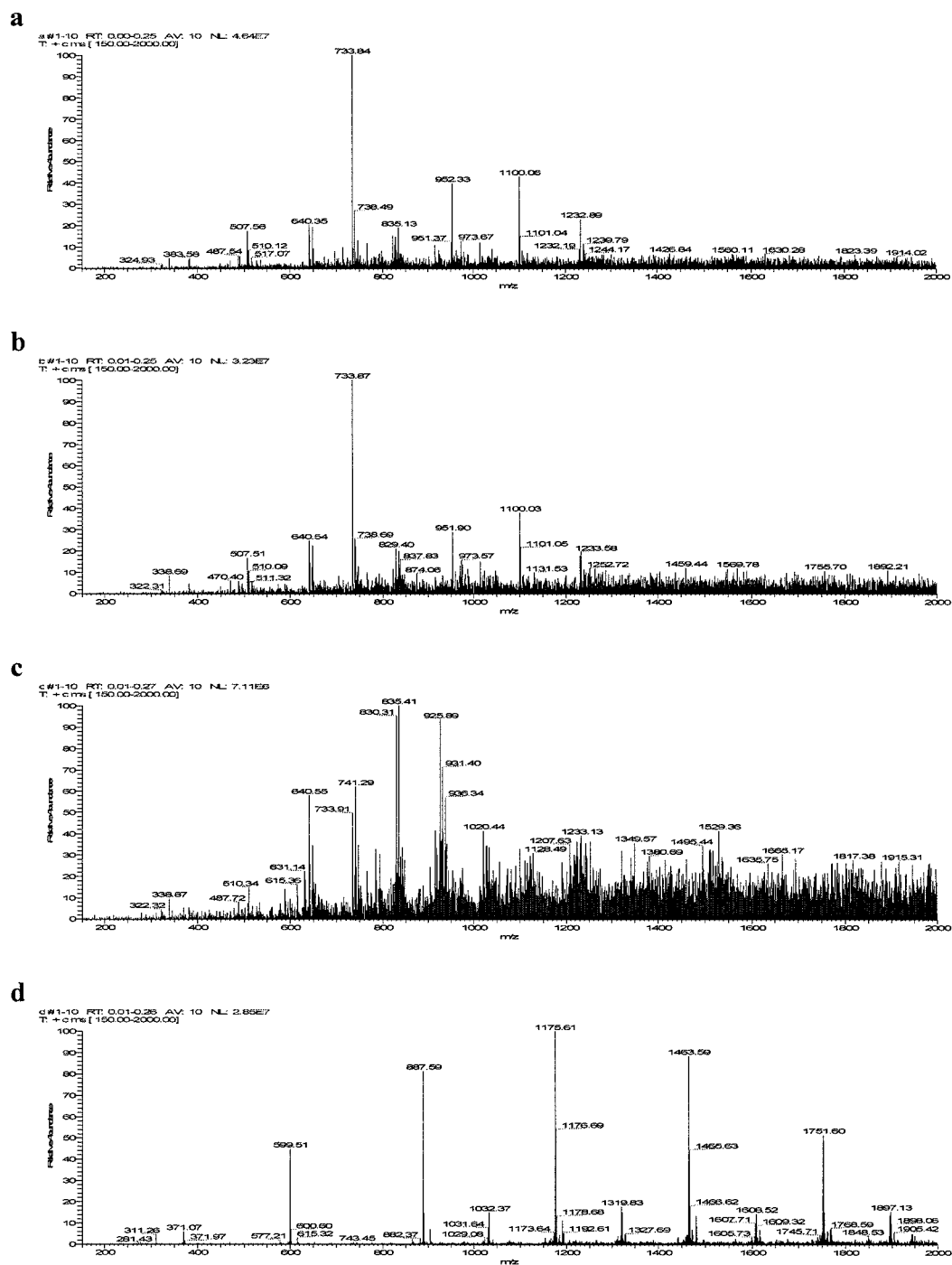
(Figure continues on next page)

E**F****G****H**

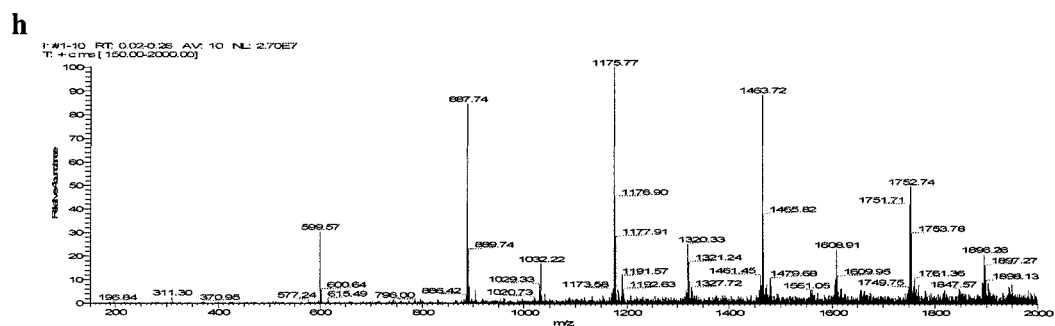
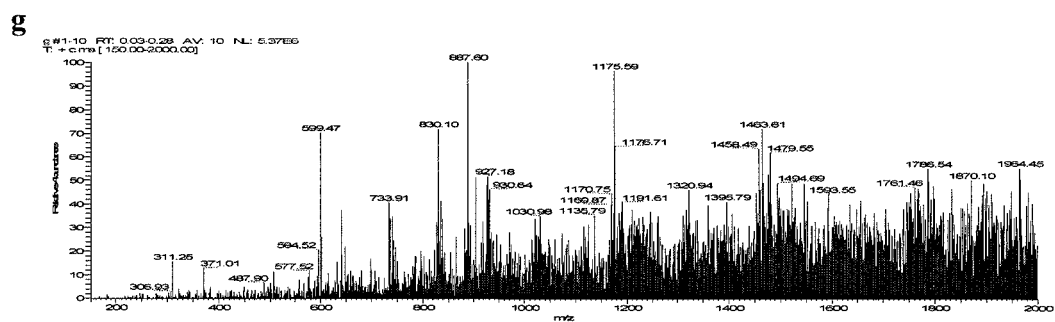
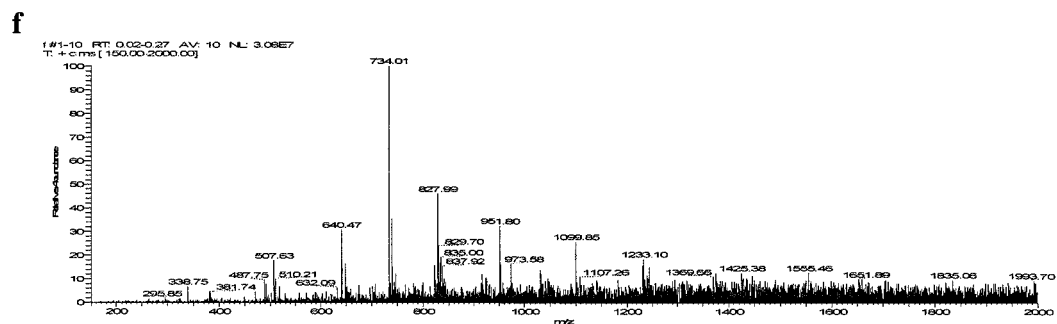
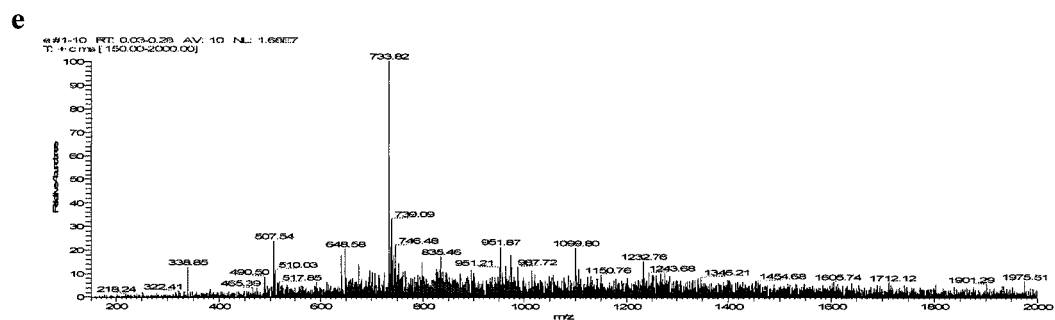
(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)

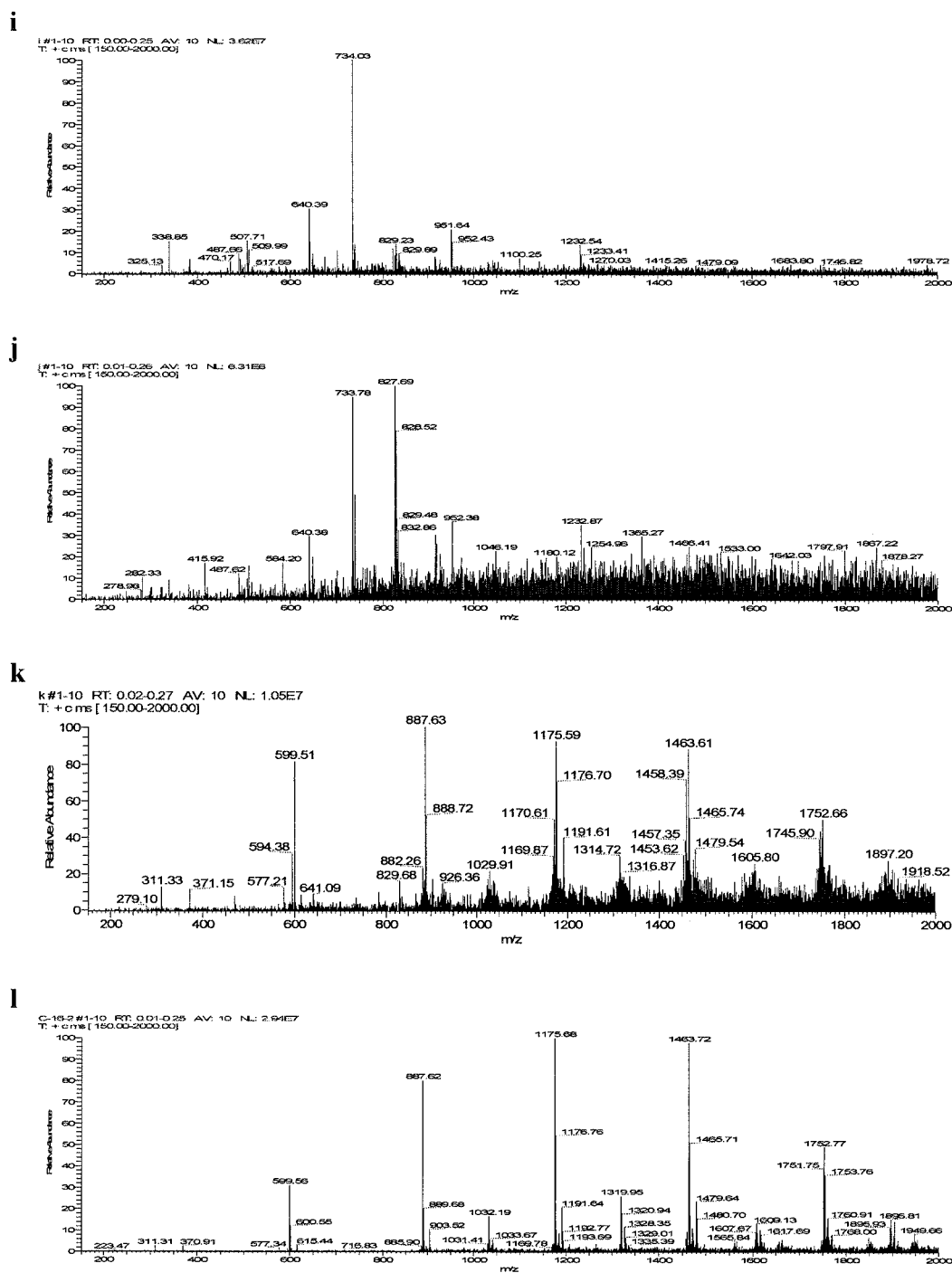
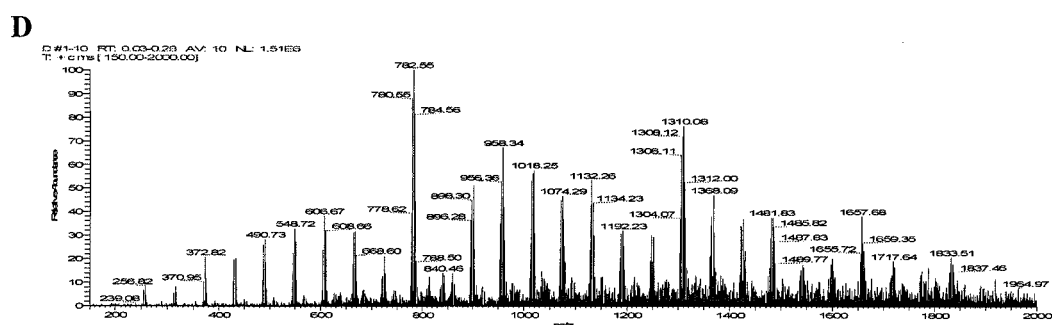
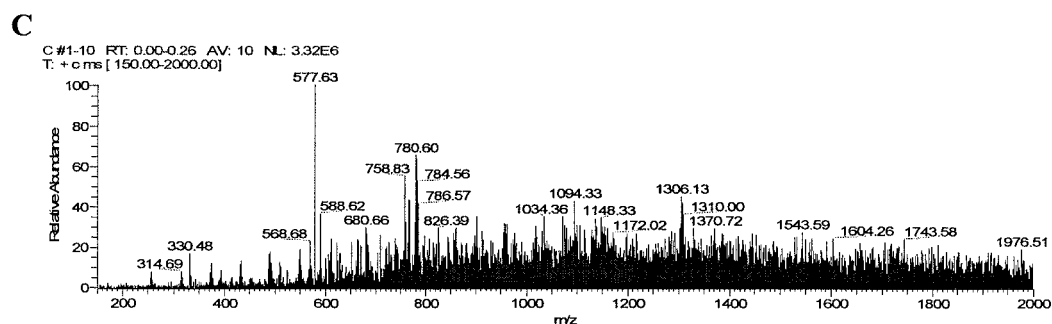
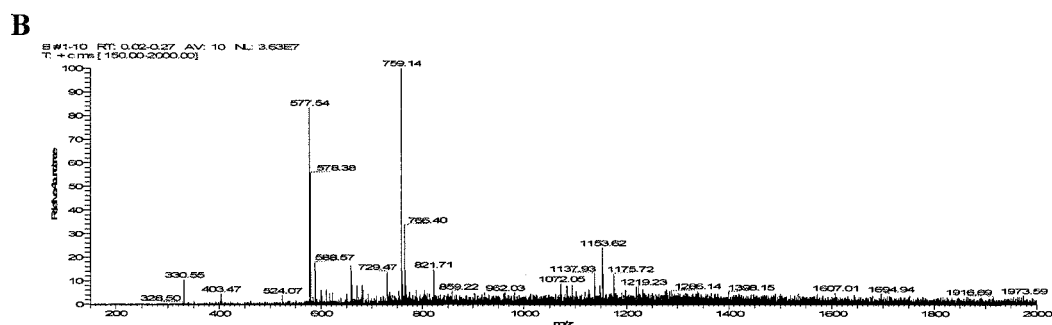
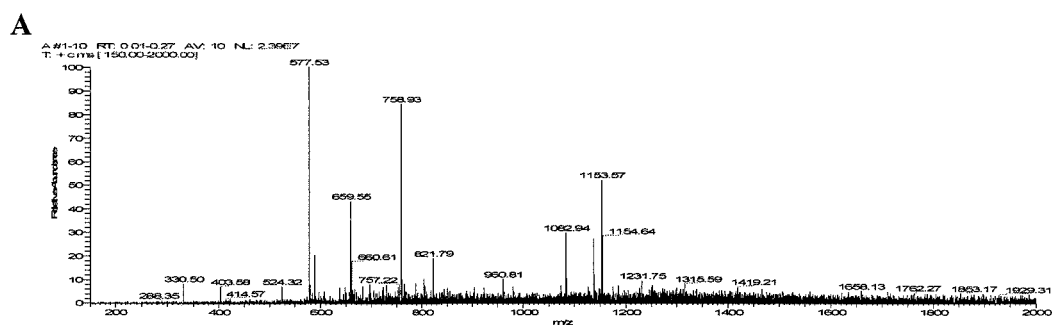
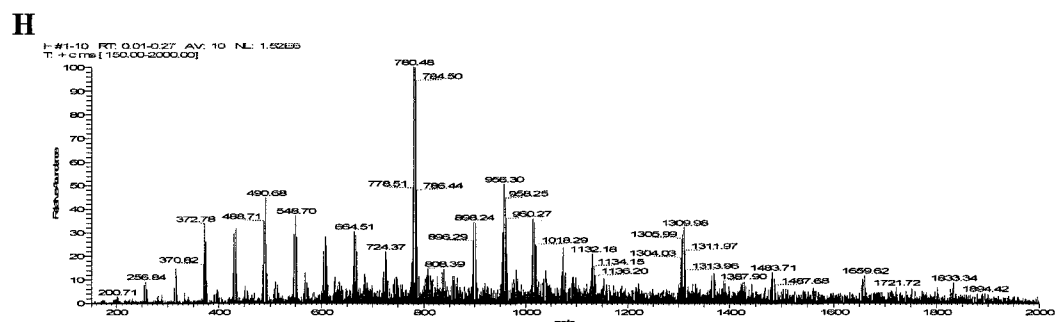
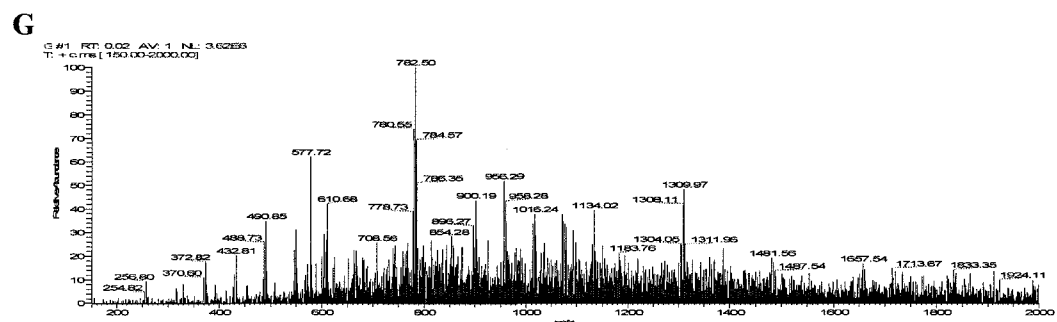
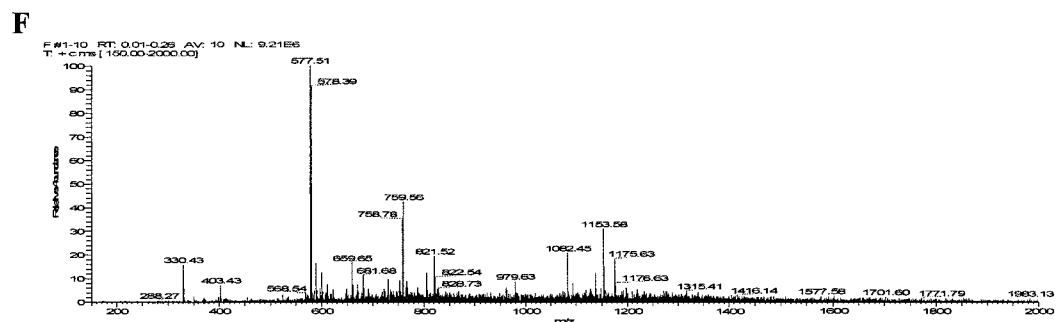
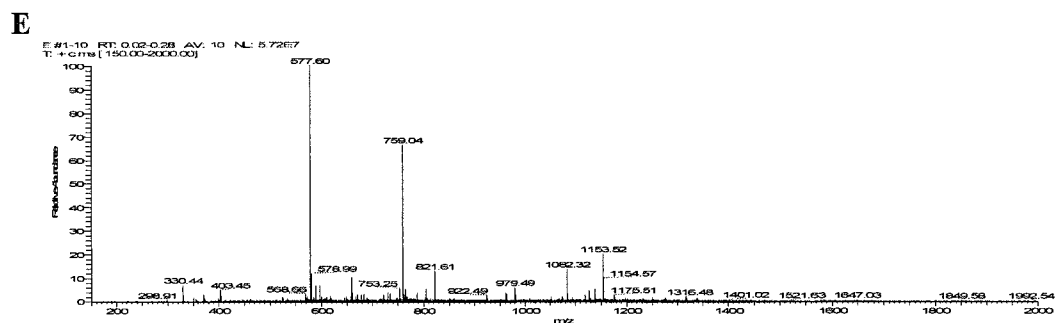


Figure F7 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using carbonic anhydrase digests.

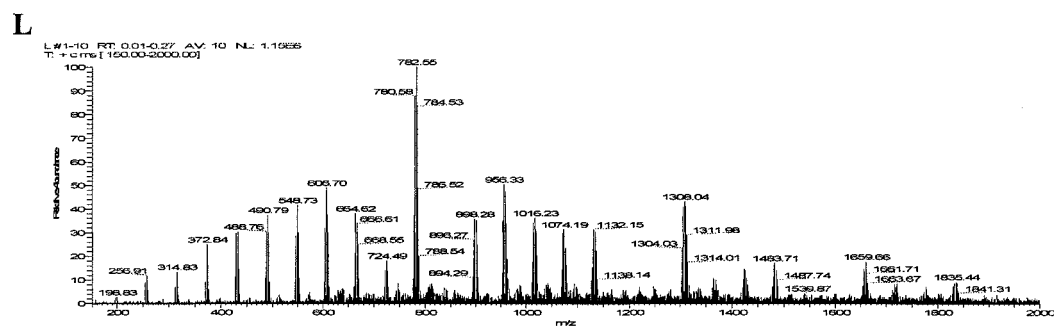
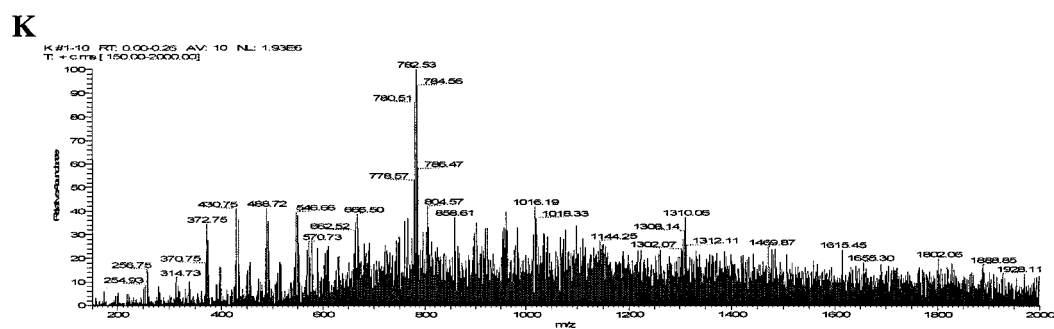
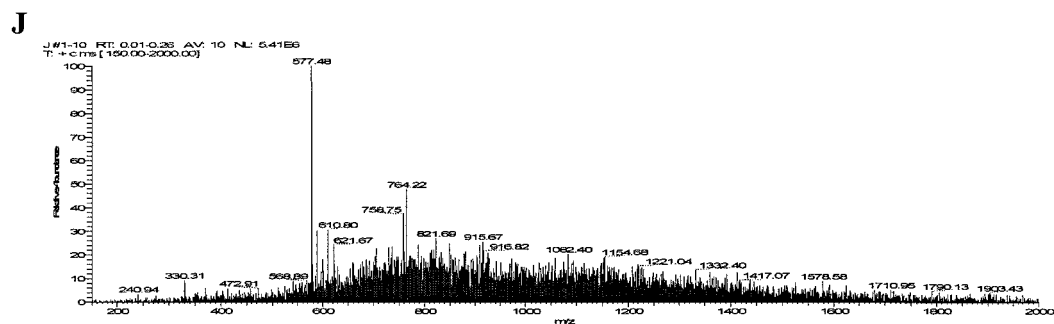
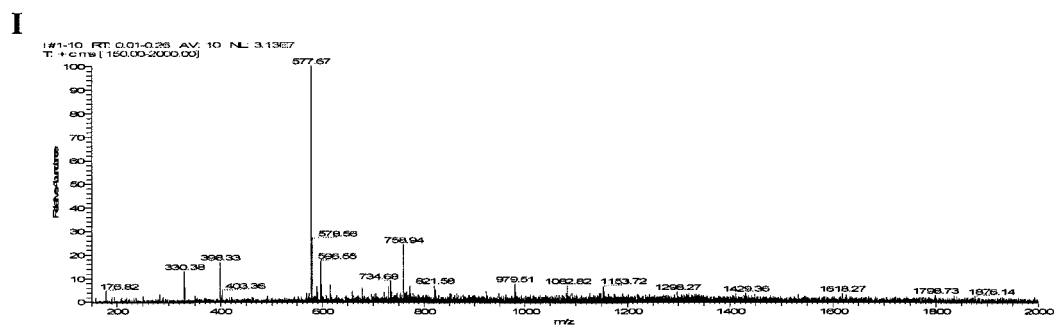
The sample conditions (A-L and a-l) are given in Table 2.10.



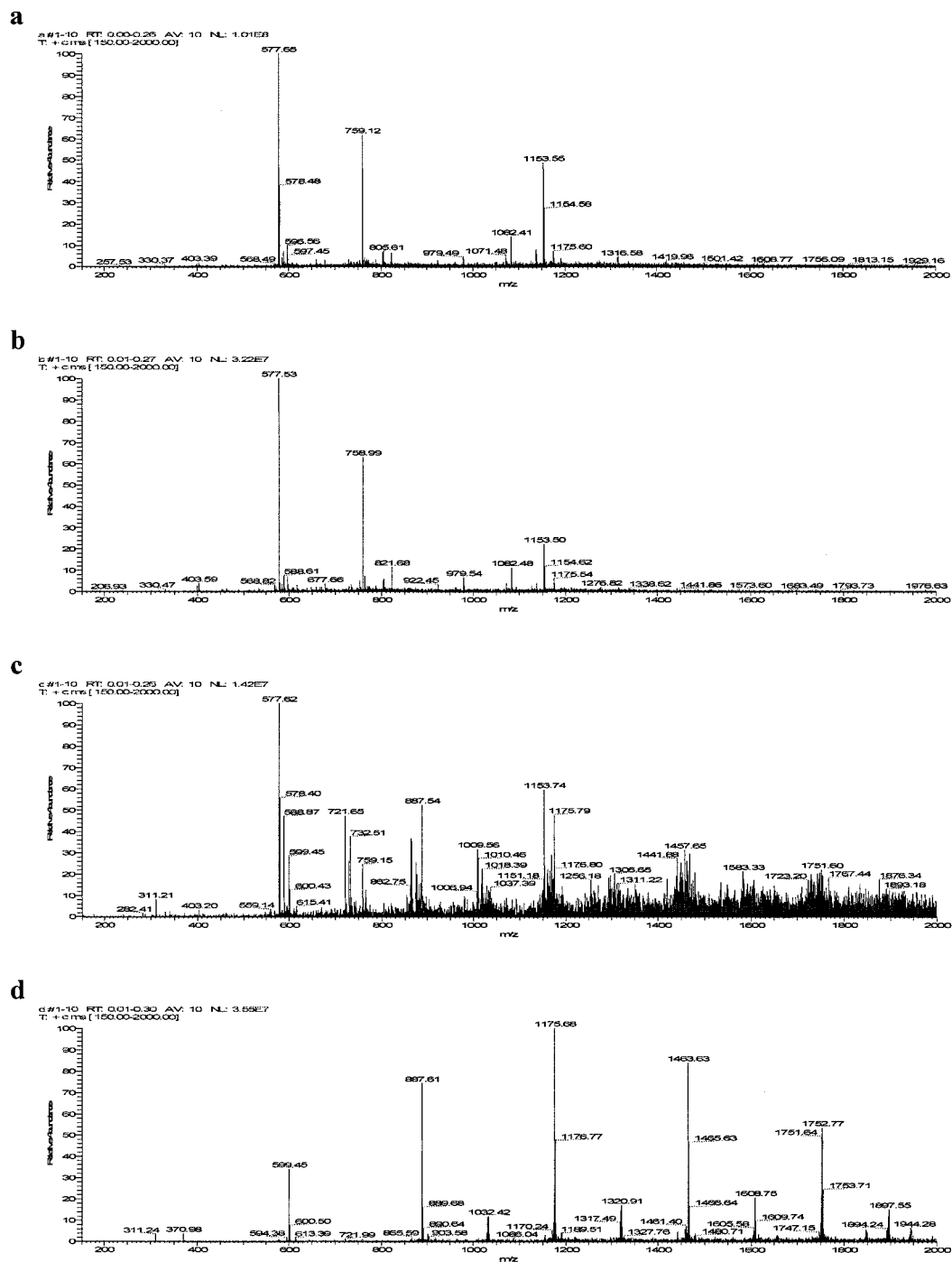
(Figure continues on next page)



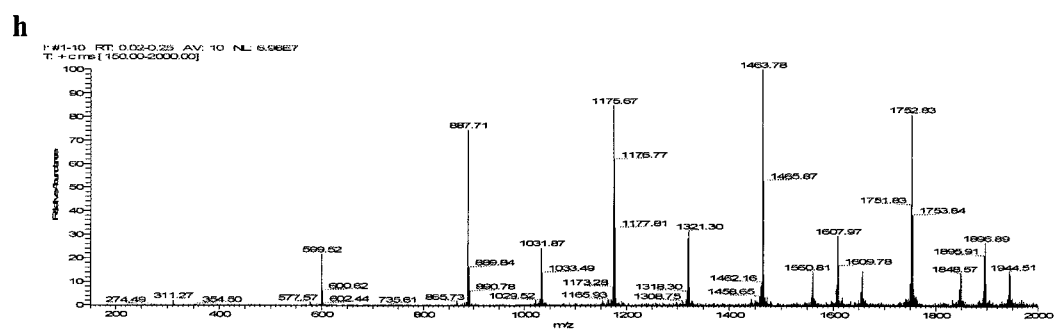
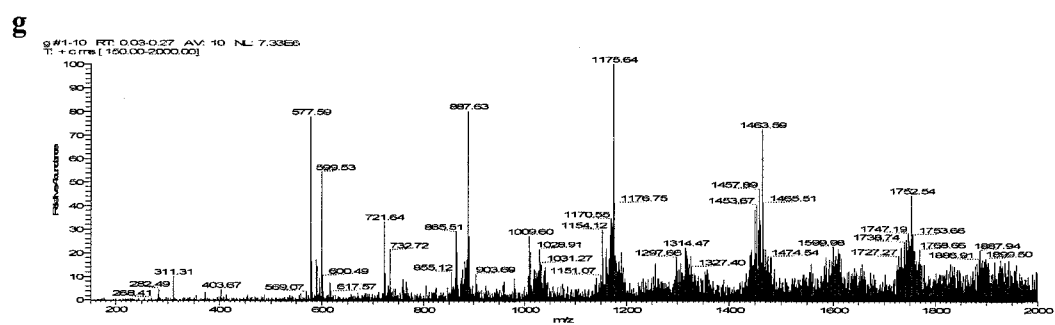
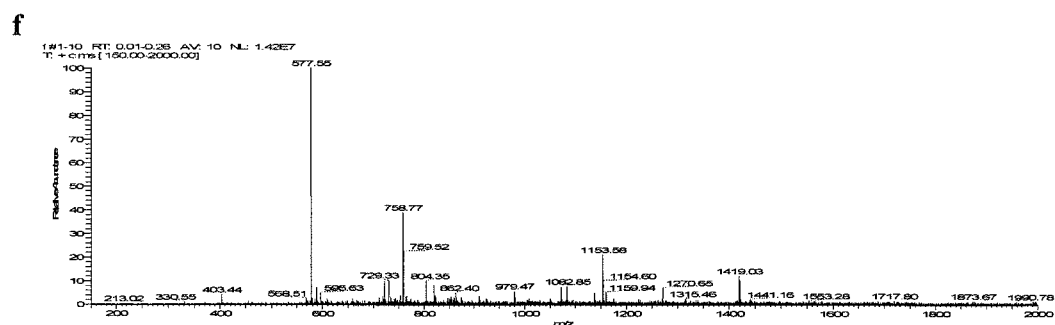
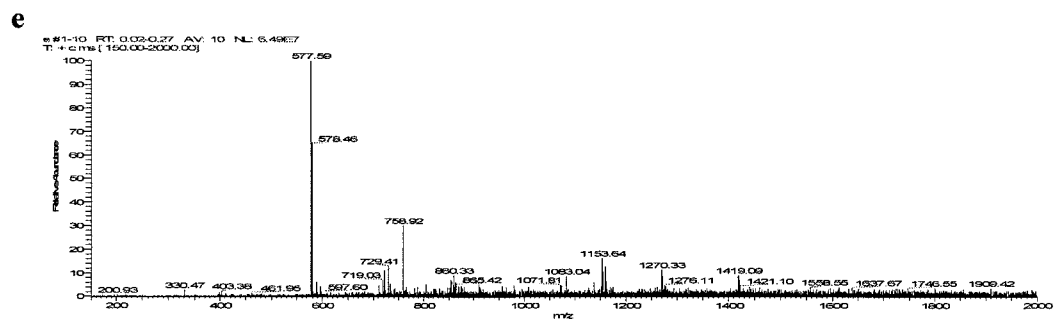
(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)



(Figure continues on next page)

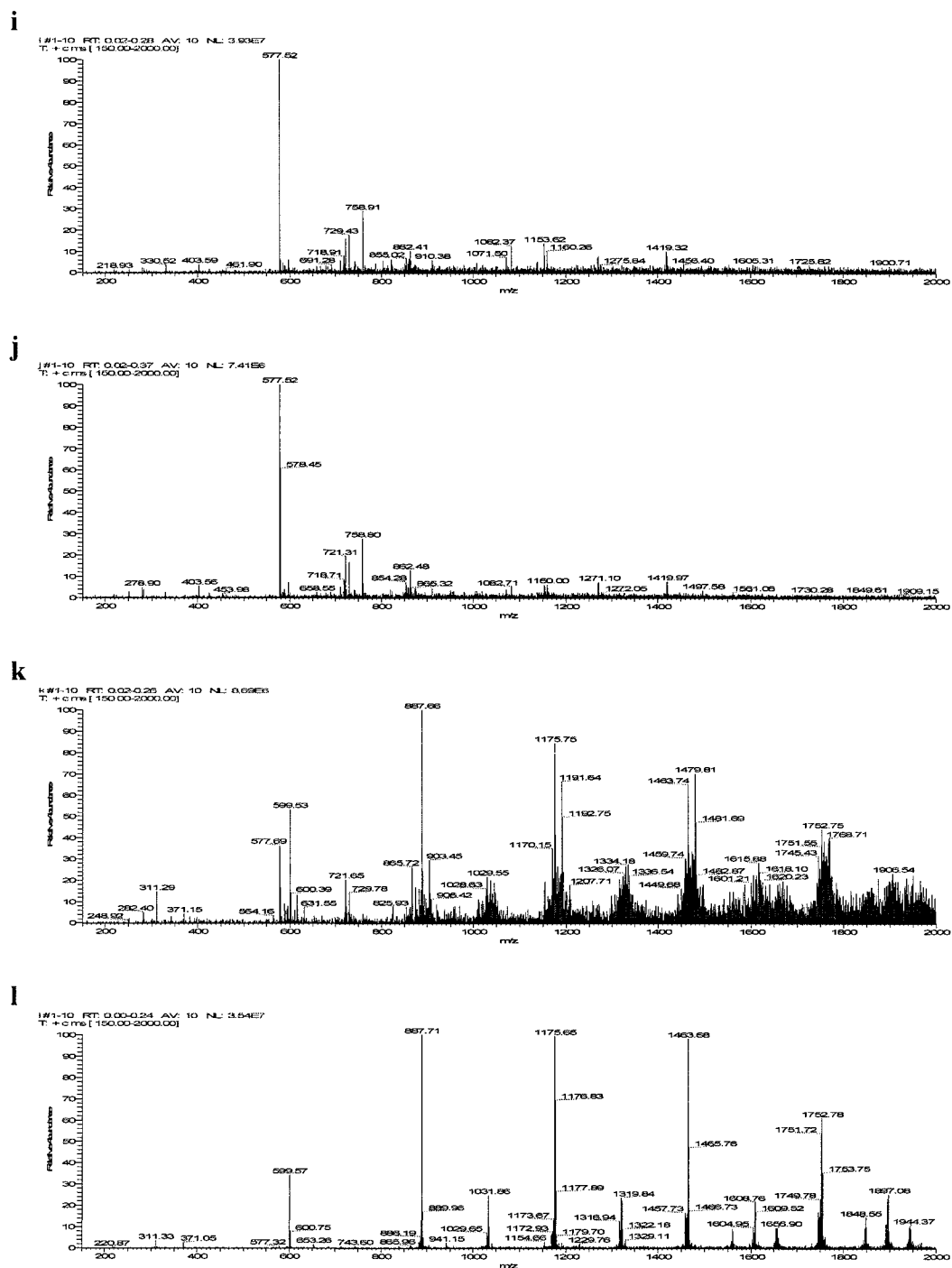


Figure F8 Analysis of salt and detergent tolerance of the static nanospray LCQ Advantage MS using trypsinogen digests.

The sample conditions (A-L and a-l) are given in Table 2.10.

Appendix G

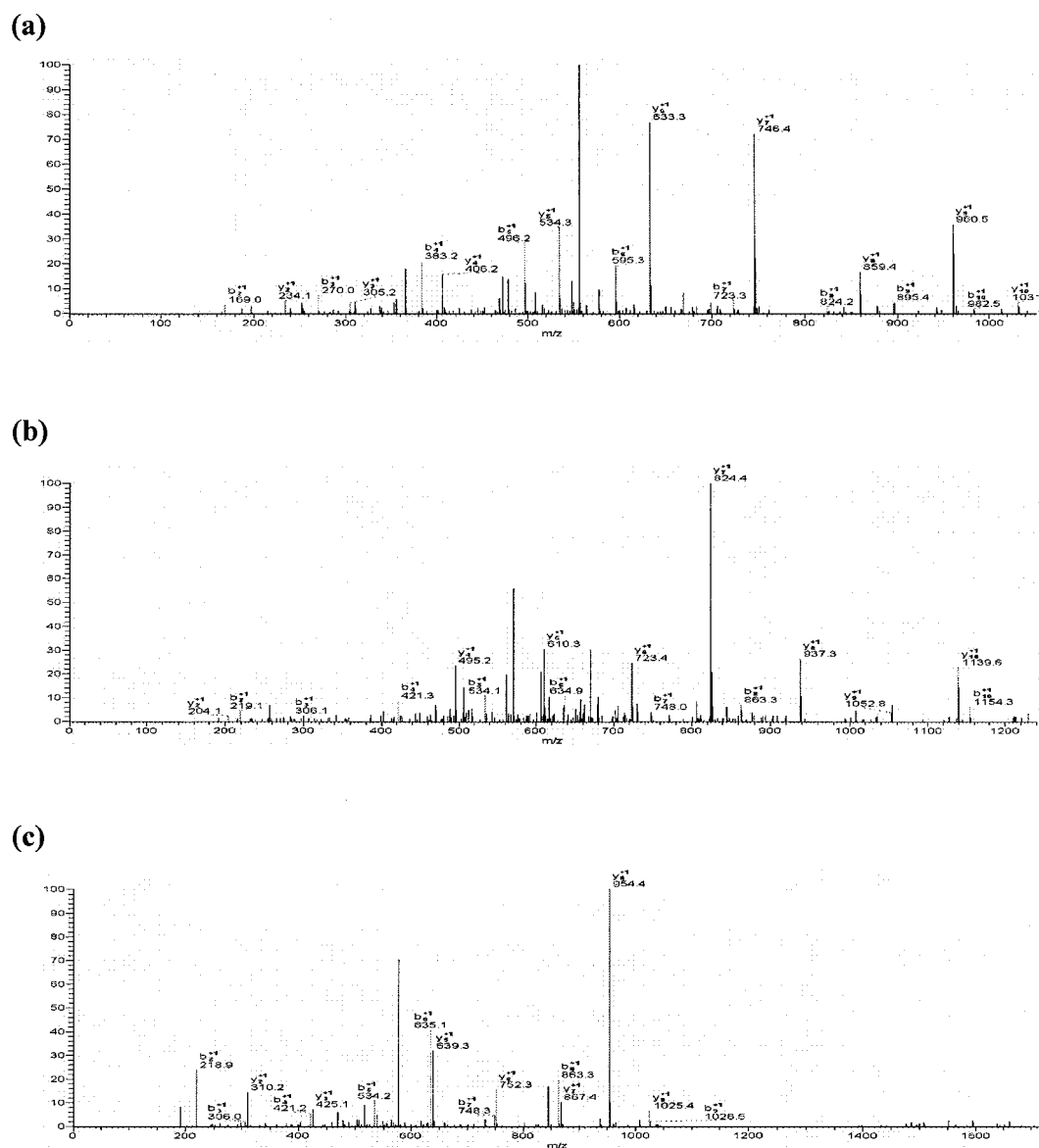


Figure G1 MS/MS spectra of identified peptides from putative phosphotransferase system phosphohistidine-containing protein

(a) Precursor ion of PATLLVQTASK with an m/z ratio of 564.98; (b) Precursor ion of FASDITLDYK GK with an m/z ratio of 679.53; (c) Precursor ion of FASDITLDYK with an m/z ratio of 587.11.

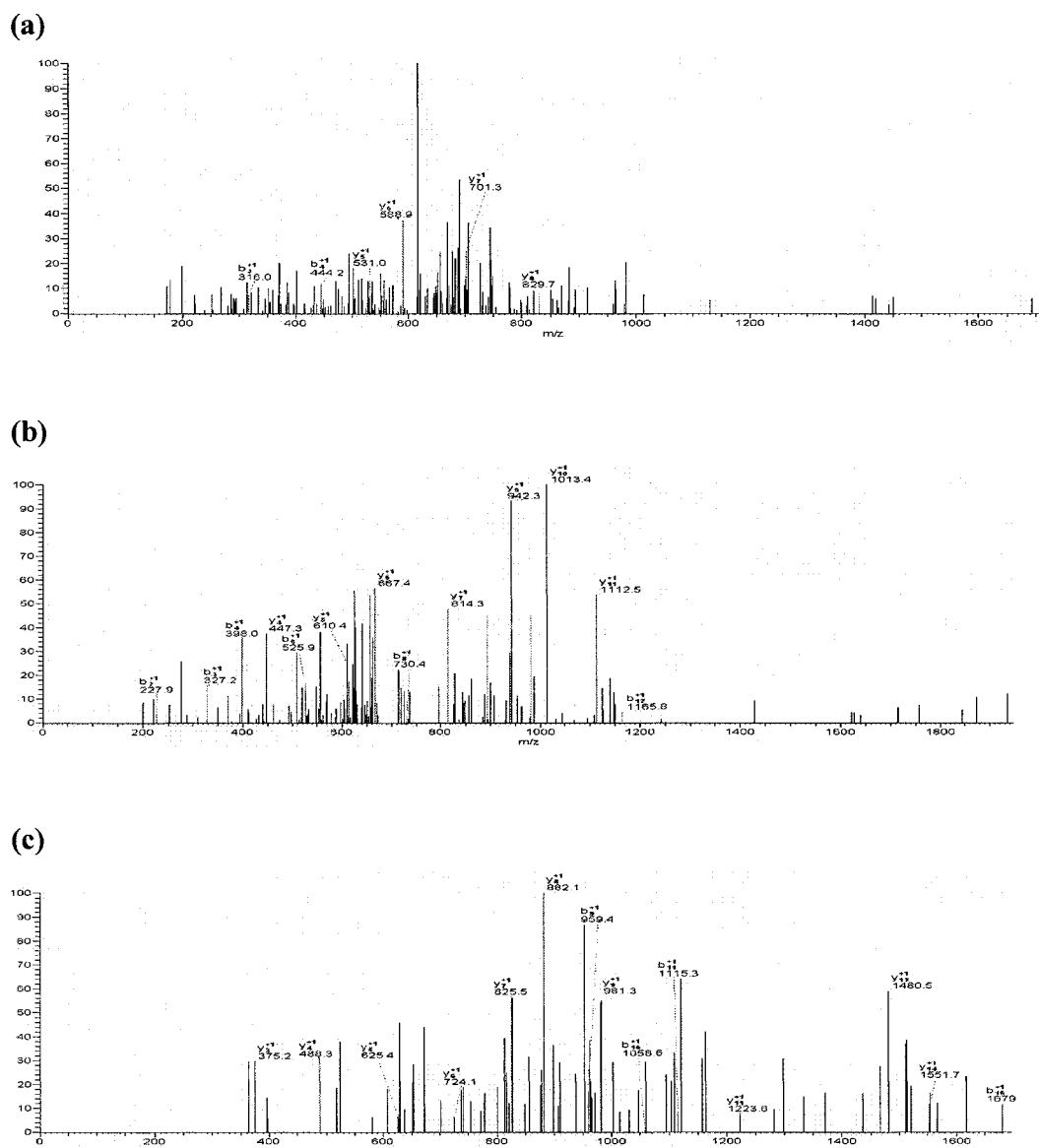


Figure G2 MS/MS spectra of identified peptides from Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase

(a) Precursor ion of DAELKIGAILSK with an m/z ratio of 573.12; (b) Precursor ion of NIVAGAFGYSGQR with an m/z ratio of 670.61; (c) Precursor ion of AFGIAEQLEVGTIVHLNNK with an m/z ratio of 970.60.

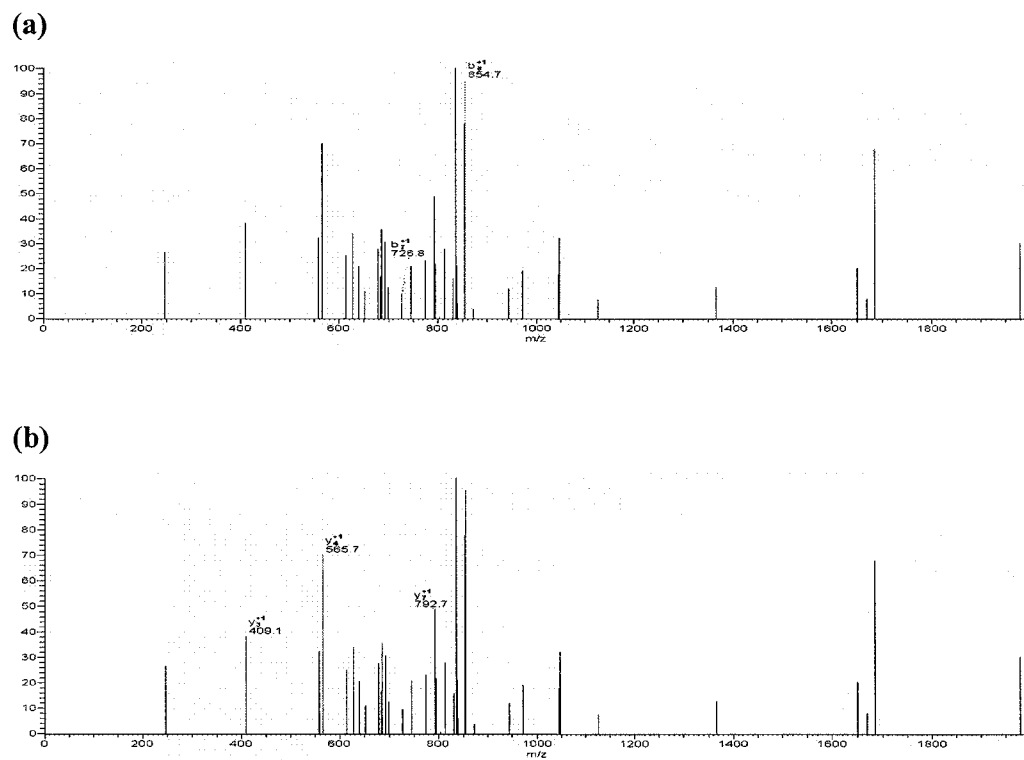


Figure G3 MS/MS spectra of identified peptides from M1 protein precursor

(a) Precursor ion of ALELAIDQASQDYNR with an m/z ratio of 854.52; (b) Precursor ion of ARLENAMEVAGRDFK with an m/z ratio of 854.52.

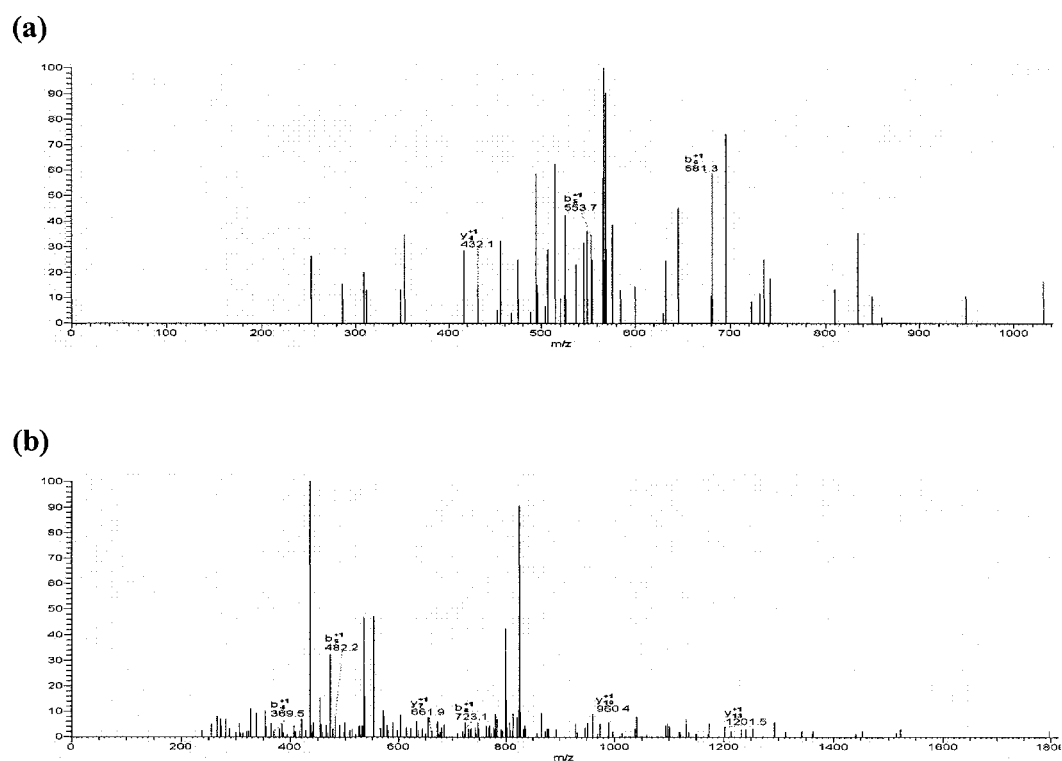


Figure G4 MS/MS spectra of identified peptides from hypothetical protein SPy1143

(a) Precursor ion of AVILRQGAITK with an m/z ratio of 585.24; (b) Precursor ion of TGPLIGPSANLSGKASGR with an m/z ratio of 842.54

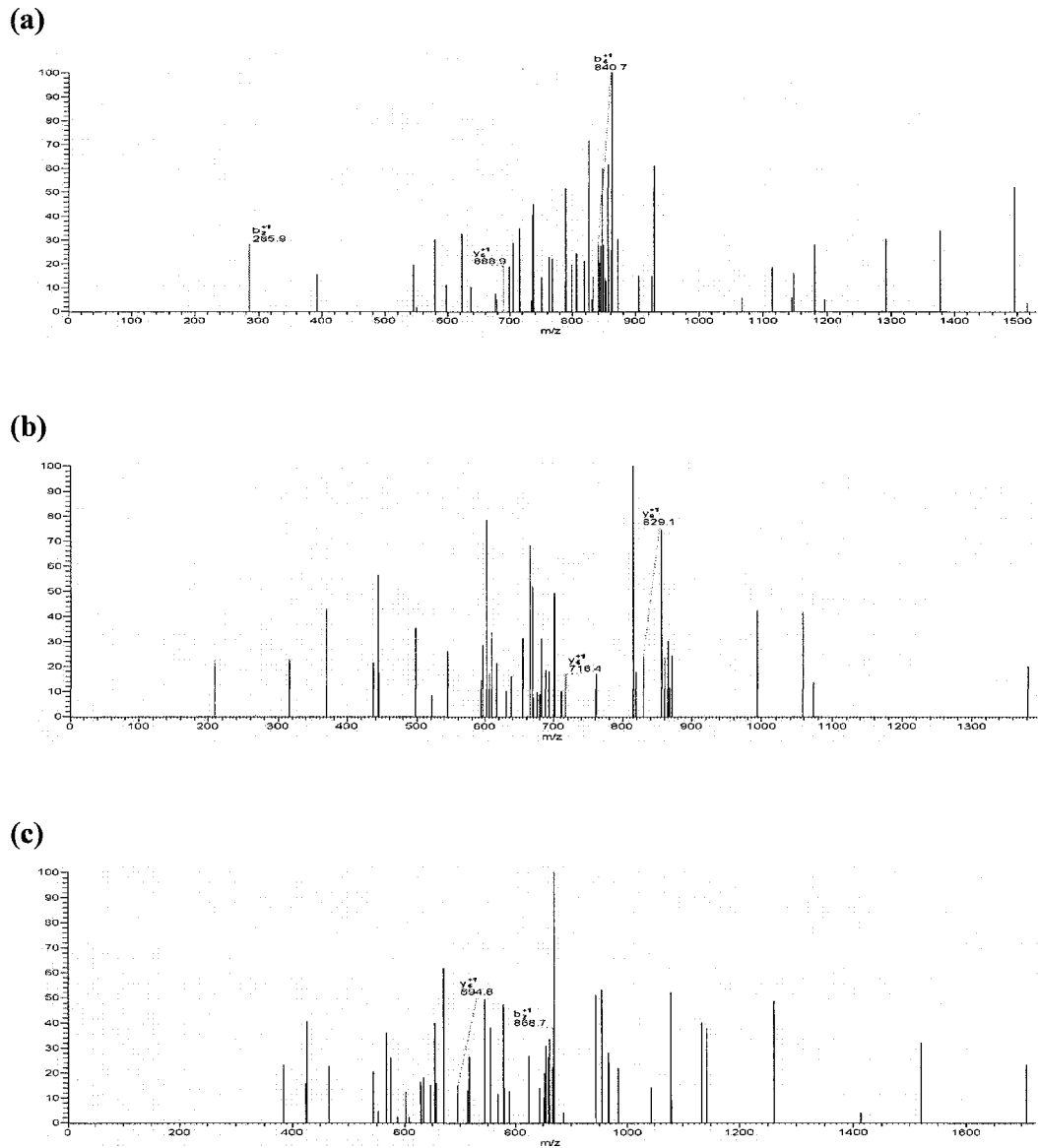
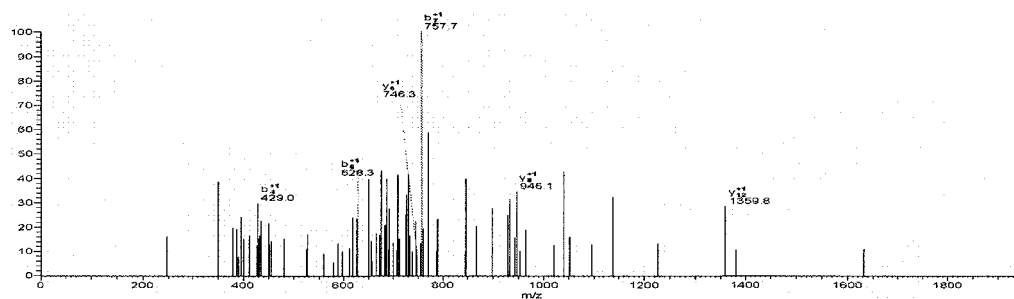


Figure G5 MS/MS spectra of identified peptides from histidine protein kinase
 (a) Precursor ion of REPFYFDDMNAIK with an m/z ratio of 880.17; (b) Precursor ion of DLVFLLRHFEK with an m/z ratio of 709.74; (c) Precursor ion of HFEKAVRNPLAHLIK with an m/z ratio of 887.18

(a)



(b)

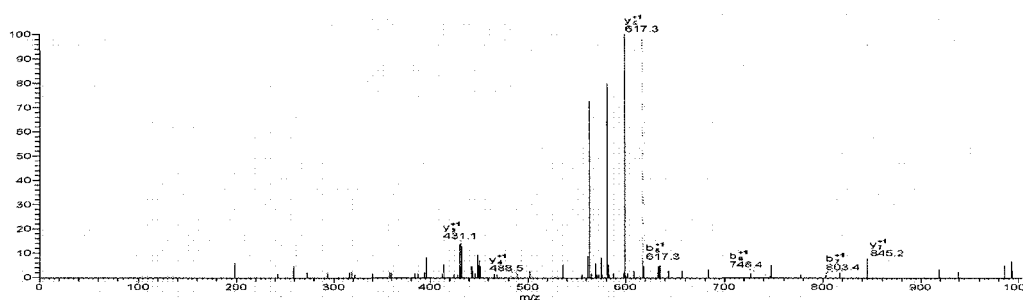
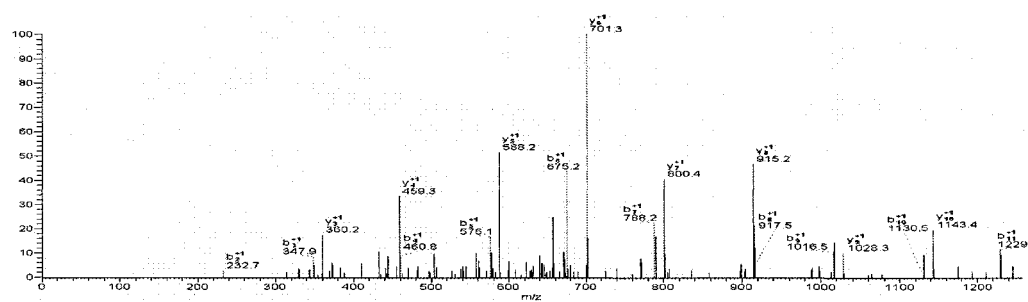


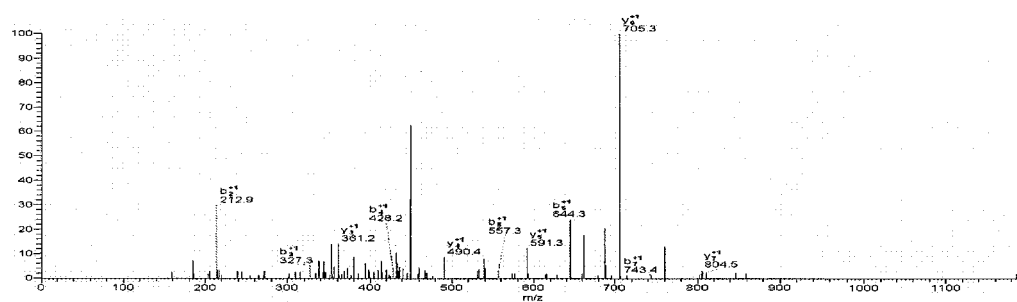
Figure G6 MS/MS spectra of identified peptides from tRNA modification GTPase

(a) Precursor ion of VDLTQAEAVMDIIR with an m/z ratio of 788.37; (b) Precursor ion of TGTIIDEVMVSVMLAPK with an m/z ratio of 617.71

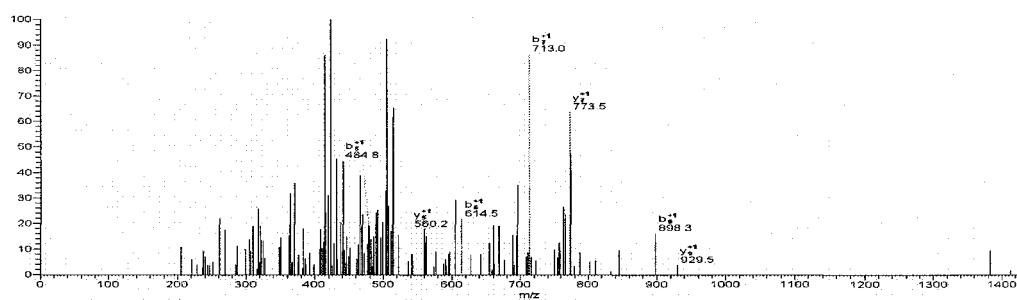
(a)



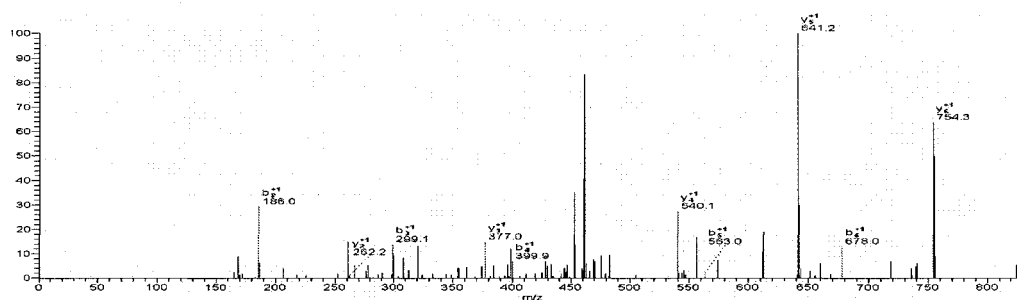
(b)



(c)

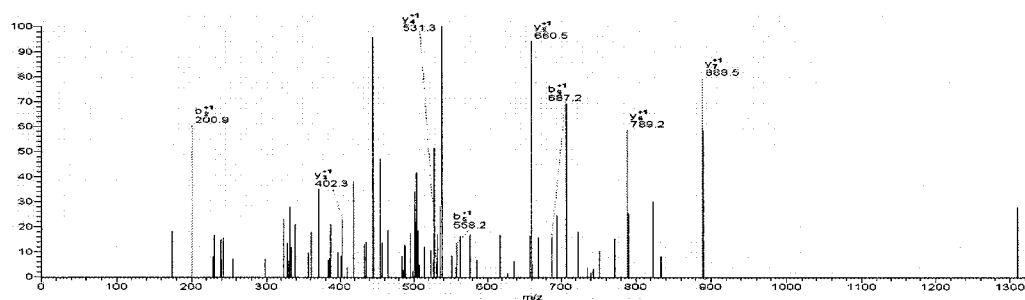


(d)

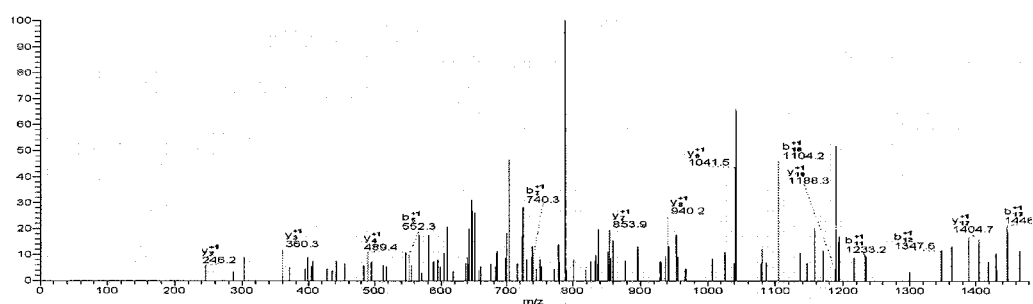


(Figure continues on next page)

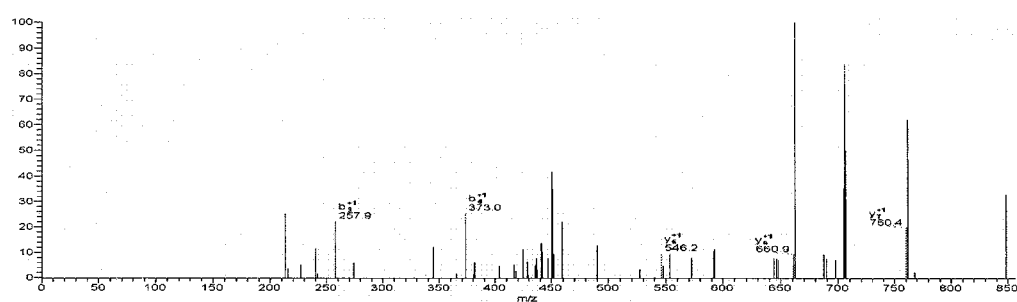
(e)



(f)



(g)



(Figure continues on next page)

(h)

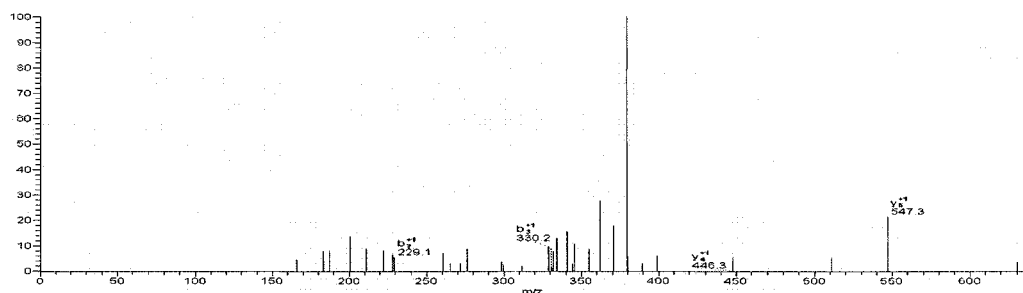


Figure G7 MS/MS spectra of identified peptides from Hypothetical protein SPy1262

(a) Precursor ion of MTDLDVIEVNVK with an m/z ratio of 688.85; (b) precursor ion of LVNTESVR with an m/z ratio of 459.53; (c) precursor ion of DGVNVEVGKK with an m/z ratio of 523.01; (d) precursor ion of GQLTYDDK with an m/z ratio of 470.83; (e) precursor ion of SIVEEEVKR with an m/z ratio of 545.01; (f) precursor ion of STSEFTSHQVENVK with an m/z ratio of 797.27; (g) precursor ion of ASVDNGVEK with an m/z ratio of 469.13; (h) precursor ion of DLTSAIR with an m/z ratio of 388.27.

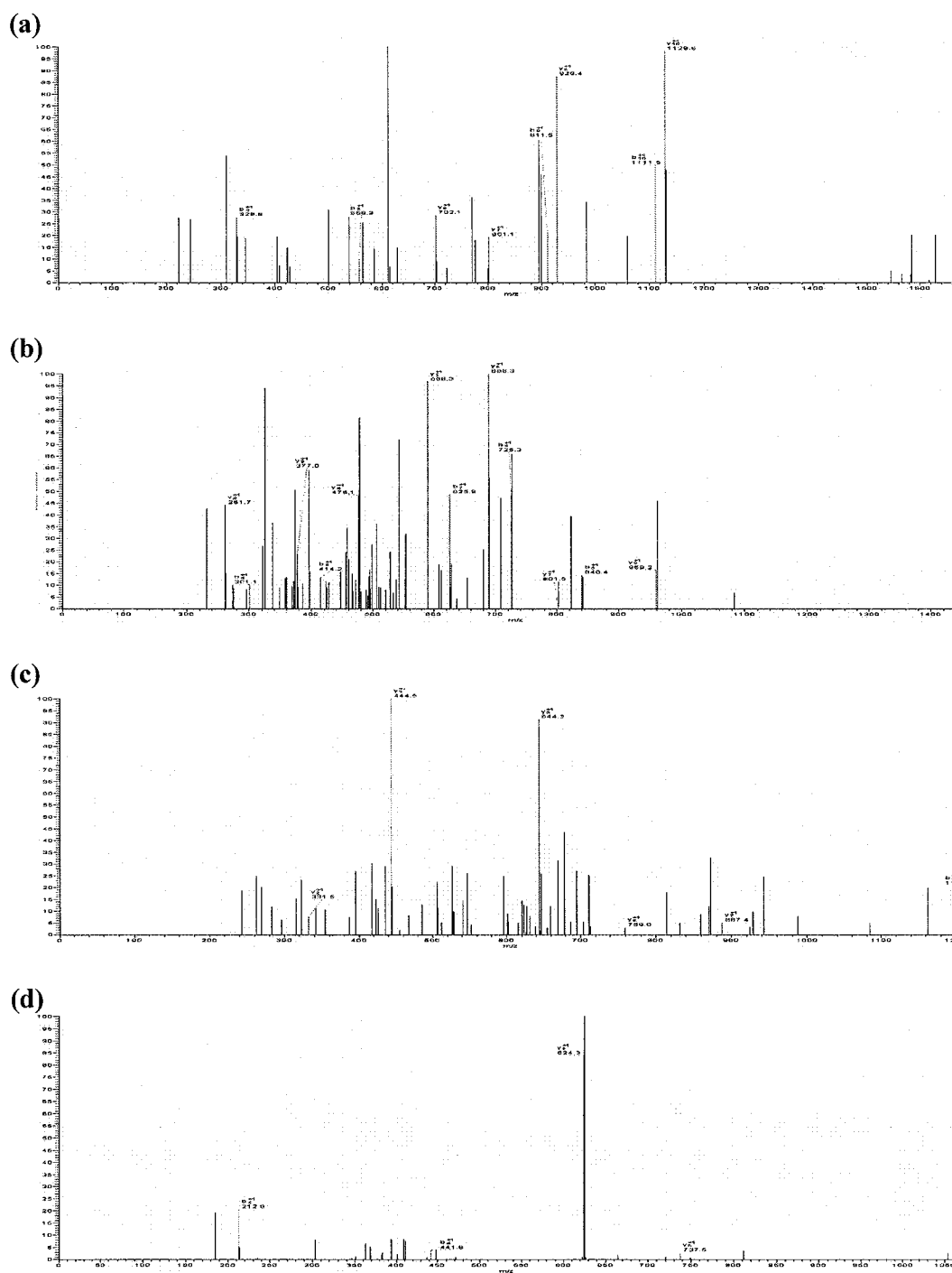
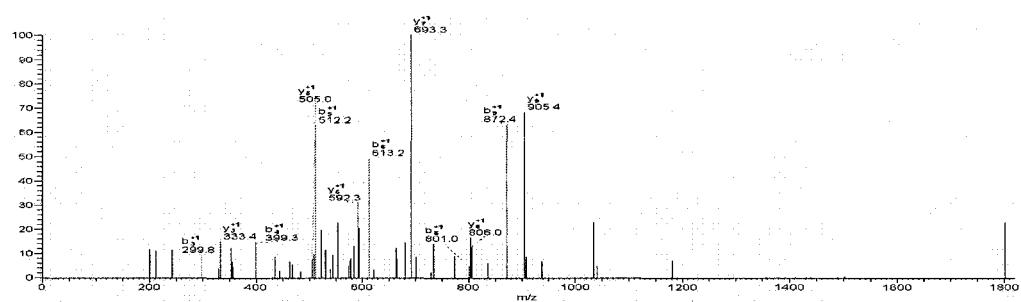


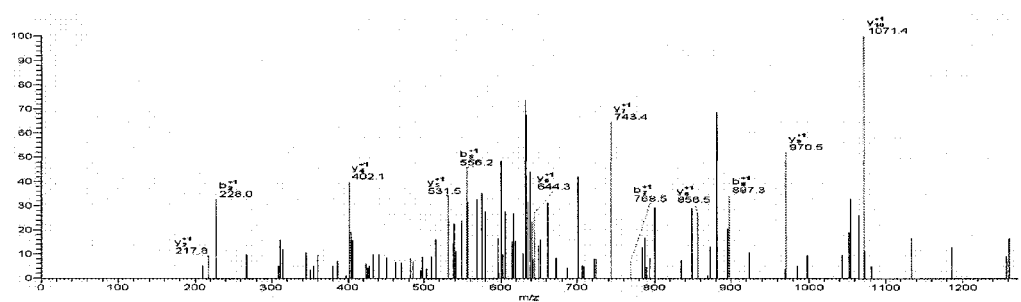
Figure G8 MS/MS spectra of identified peptides from 50S ribosomal protein L10.

(a) Precursor ion of KAEQVELIAEK with an m/z ratio of 629.48; (b) precursor ion of AAASIVIVDSR with an m/z ratio of 552.02; (c) precursor ion of GLTVDQDTVLRR with an m/z ratio of 687.30; (d) precursor ion of VINDFTK with an m/z ratio of 419.15;

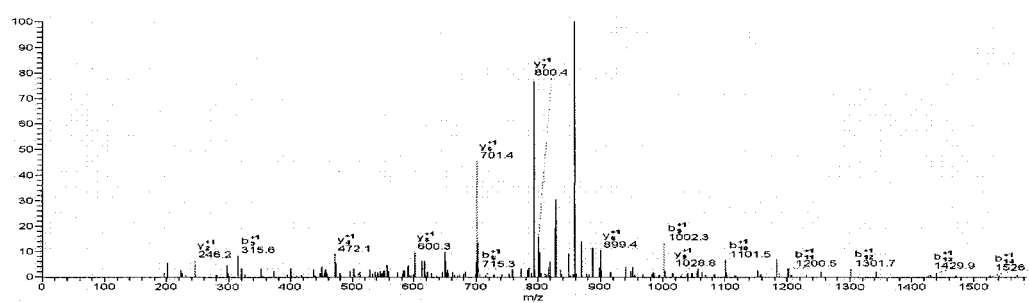
(a)



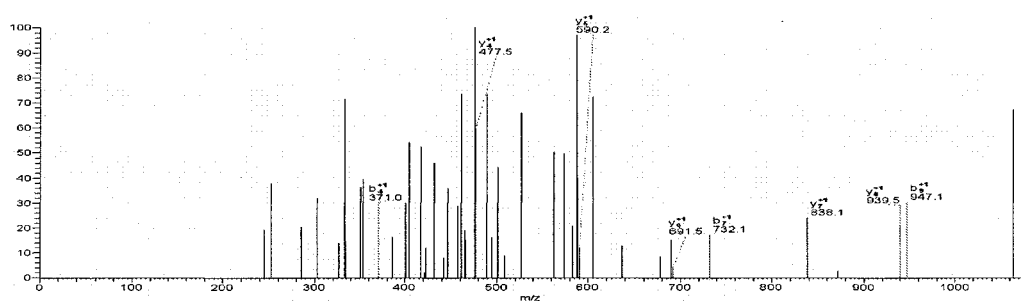
(b)



(c)



(d)



(Figure continues on next page)

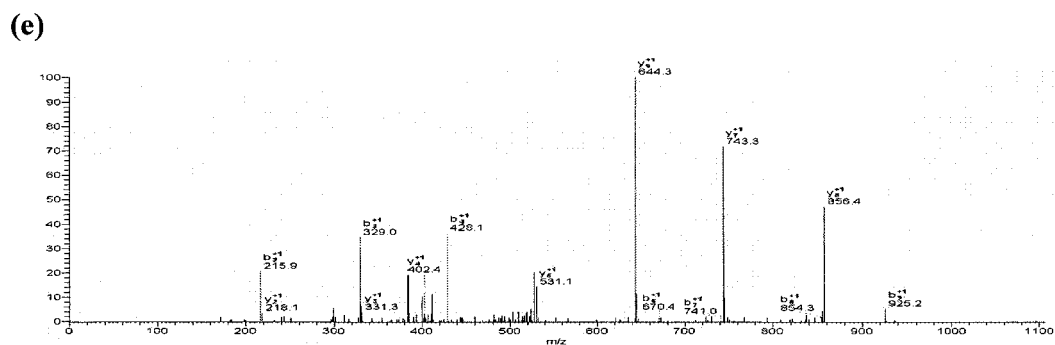


Figure G9 MS/MS spectra of identified peptides from RopA

(a) Precursor ion of AVEVITSTASVK with an m/z ratio of 603.09; (b) precursor ion of VKTNLVIEAIAK with an m/z ratio of 650.38; (c) precursor ion of GKEWTLAEVVTKEPVK with an m/z ratio of 634.84; (d) precursor ion of GVITFTISQDK with an m/z ratio of 605.08; (e) precursor ion of TNLVIEAIAK with an m/z ratio of 536.07.

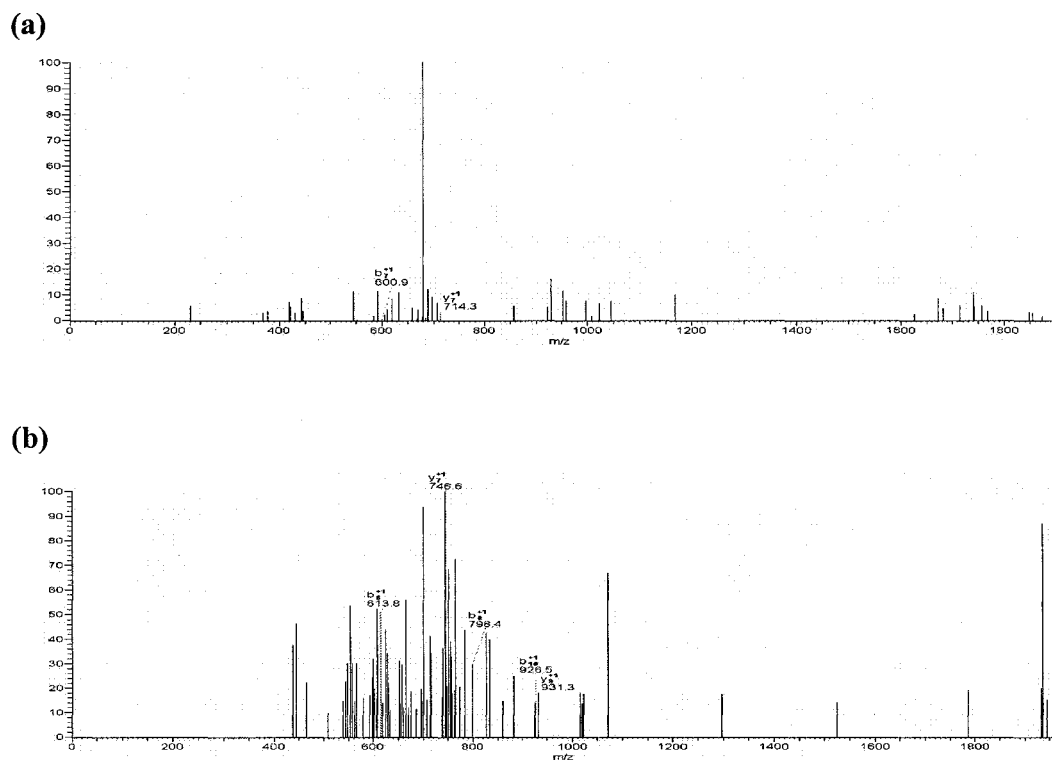


Figure G10 MS/MS spectra of identified peptides from Collagen-like surface protein

(a) Precursor ion of GETFAQGPVGPQGEK with an m/z ratio of 706.29; (b) precursor ion of GEQGIQ GKAGEKGER with an m/z ratio of 773.20.

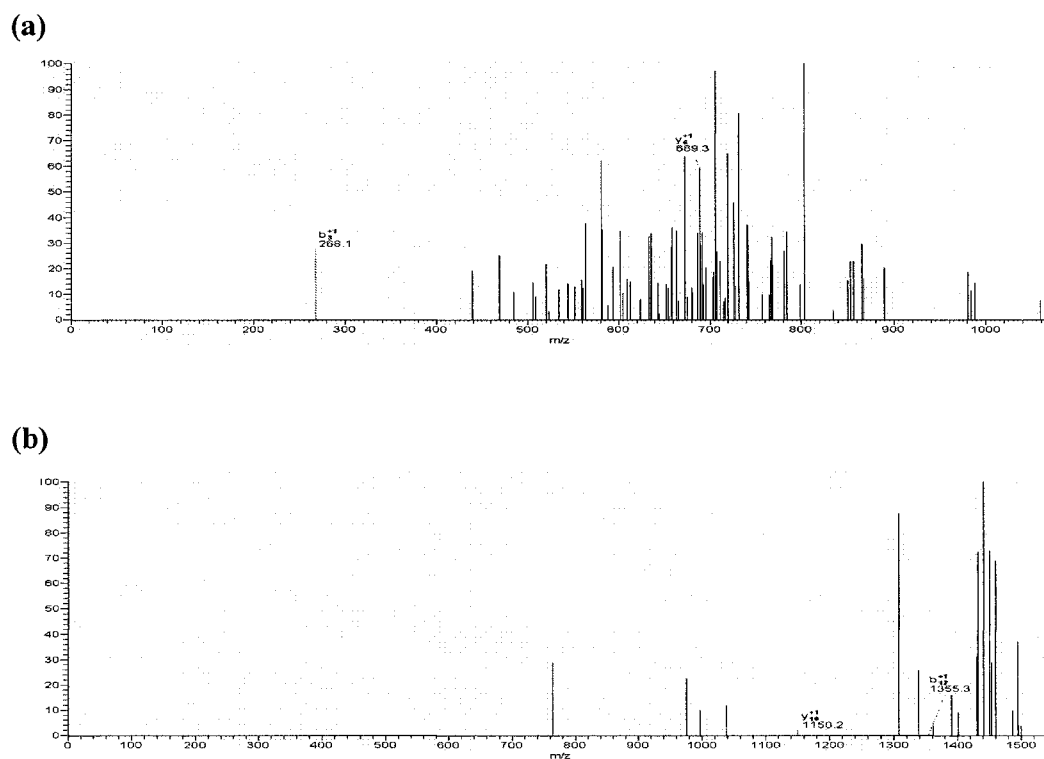


Figure G11 MS/MS spectra of identified peptides from Putative carbamoyl-phosphates synthase small subunit

(a) Precursor ion of GIPGISGIDTRALTK with an m/z ratio of 749.98; (b) precursor ion of MTFGHRGFNHAVR with an m/z ratio of 1529.94.

Published paper

RESEARCH ARTICLE

Group A streptococcus cell-associated pathogenic proteins as revealed by growth in hyaluronic acid-enriched media

Meng Zhang¹, Fiona M. McDonald², Shane S. Sturrock², Simon J. Charnock^{1*}, Ian Humphery-Smith² and Gary W. Black^{1**}

¹ Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK

² Biosystems Informatics Institute, Bioscience Centre, International Centre for Life, Newcastle upon Tyne, UK

Group A streptococcus (GAS), also known as *Streptococcus pyogenes*, is a human pathogen and can cause several fatal invasive diseases such as necrotising fasciitis, the so-called flesh-eating disease, and toxic shock syndrome. The destruction of connective tissue and the hyaluronic acid therein, is a key element of GAS pathogenesis. We therefore propagated GAS in hyaluronic acid (HA)-enriched growth media in an attempt to create a simple biological system that could reflect some elements of GAS pathogenesis. Our results show that several recognised virulence factors were up-regulated in HA-enriched media, including the M1 protein, a collagen-like surface protein and the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase, which has been shown to play important roles in streptococcal pathogenesis. Interestingly, two hypothetical proteins of unknown function were also up-regulated and detailed bioinformatics analysis showed that at least one of these hypothetical proteins is likely to be involved in pathogenesis. It was therefore concluded that this simple biological system provided a valuable tool for the identification of potential GAS virulence factors.

Received: July 31, 2006
Revised: January 16, 2007
Accepted: January 24, 2007

Keywords:

Group A streptococcus / Hyaluronic acid-enriched growth media / Pathogenesis / Proteomic analysis / Signature peptide search

1 Introduction

Group A streptococcus (GAS) is a Gram-positive human pathogen. This bacterium is specific to humans and human disease. It is estimated that between 5–15% of normal

individuals harbour the bacteria, usually in the respiratory tract, without signs of disease [1]. However, when the bacteria are able to penetrate vulnerable tissues and when the host defences are compromised, a variety of types of suppurative infection can occur. Commonly, GAS is responsible for pharyngitis, scarlet fever, impetigo or cellulitis. Myositis, necrotizing fasciitis and streptococcal toxic shock syndrome are caused by invasive infections. It has been recently estimated that severe GAS infections affect more than 18 million people *per* year with more than 0.5 million deaths [2].

Correspondence: Dr. Meng Zhang, School of Applied Sciences, Northumbria University, Ellison Building, Newcastle upon Tyne NE1 8ST, UK

E-mail: meng.zhang@northumbria.ac.uk

Fax: +44-19-1227-3519

Abbreviations: ΔC_r , difference in normalised cross correlation score; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GAS, group A streptococcus; HA, hyaluronic acid; O/E, observed/expected; SpeB, streptococcal pyrogenic exotoxin B; X_{corr} cross correlation score

* Present address: Megazyme International Ireland Limited, Bray, County Wicklow, Republic of Ireland.

** Additional corresponding author: Dr. Gary W. Black
E-mail: gary.black@northumbria.ac.uk

The degradation of connective tissue by GAS is a key element of invasive infection [3]. Hyaluronic acid (HA) is a major component of the extracellular matrix of connective tissue and is a negatively charged high molecular mass polysaccharide ($\sim M_r$ of 1 000 000) consisting of alternating residues of *N*-acetylglucosamine and glucuronic acid. It is produced by all mammals and is distributed in various tissues, with the highest concentrations found in soft connective tissues [4]. To invade the connective tissue of the human host, GAS produces hyaluronate lyases that depolymerise HA [5]. The result of the enzymatic depolymerisation of HA is a decrease in the viscosity of the extracellular matrix, and therefore increased permeability of the connective tissues, and potentially an increase in pathogen aggression, and diffusion of its toxins, through the connective tissue [6]. In addition to hyaluronate lyases, many other factors are known to contribute to streptococcal virulence, such as M proteins, lipoteichoic acid, the HA capsule that surrounds the pathogen, pyrogenic exotoxins, streptolysins and C5a peptidase [7]. All these virulence factors play a role in streptococcal infection and disease. M proteins, lipoteichoic acid and C5a peptidase are required for successful adherence and protection of GAS from constitutive host defences. Streptolysins and pyrogenic exotoxins function to destroy tissues as toxins and proteases.

Despite the large number of studies that have analysed the expression of the GAS virulence factors, the molecular basis of many GAS infections remains largely unknown. Classical approaches, such as the generation and characterization of isogenic mutants, are difficult to apply to an in-depth analysis of the vast number of bacterial gene products potentially involved in virulence [8]. Therefore, to examine simultaneous changes in multiple virulence factors, proteomic analysis of GAS is more appropriate. For instance, Lei *et al.* [9] compared the culture supernatants of wild-type GAS strains and isogenic mutant strains and identified previously undescribed extracellular proteins, including several that are immunogenic in the course of host–pathogen interactions. Thongboonkerd *et al.* [10] showed that fluoride caused the decreased expression of proteins used to respond to stress, virulence factors, and proteins implicated in several GAS non-suppurative diseases. Chaussee *et al.* [11] compared proteome maps of a GAS wild-type strain and a *rgg* mutant strain and concluded that *Rgg* is a global regulatory factor that contributes to growth phase-dependent synthesis of proteins associated with secondary metabolism and oxidative and thermal stress responses. In addition, proteome analysis has been used to confirm findings from transcriptomics experiments and to distinguish between transcriptional and post-transcriptional regulatory events [8].

Our approach was to grow GAS in growth media enriched in HA, in an attempt to create a simple biological system that could reflect some elements of GAS pathogenesis. High-resolution separation of the resulting proteins by 2-DE, in-gel trypsin digestion and static nanospray IT MS were used to characterise proteins expressed by the patho-

gen. Further functional analysis of these proteins was performed using the signature peptide search tool, TSSview™. We show that several proteins involved in pathogenesis were differentially expressed and two hypothetical proteins of unknown function were up-regulated. Interestingly, detailed bioinformatics analysis showed that at least one hypothetical protein is likely to be involved in GAS pathogenesis.

2 Materials and methods

2.1 Bacteria and growth conditions

The M1 serotype of *Streptococcus pyogenes*, strain SF370 (ATCC 700294 [12]), was employed in this study. The bacteria were kept on Columbia Agar supplemented with 5% v/v horse blood, and were grown either in Brain Heart Infusion broth (Oxoid, UK) supplemented with 0.5% w/v sodium hyaluronate (Fisher Scientific, UK), referred to hereafter as HA-enriched media, or in unsupplemented Brain Heart Infusion broth, referred to hereafter as control media. Incubations were performed at 37°C overnight until the culture reached stationary phase, *i.e.* an optical density at 660 nm between 1.0 and 1.2.

2.2 Extraction of proteins

To isolate protein extracts from strain SF370, 80 mL of culture was harvest by centrifugation at $4000 \times g$ at 4°C for 30 min. The culture supernatant was then removed, and the cell pellet resuspended and washed in 5 mL of ice-cold PBS (140 mM NaCl, 10 mM Na₃PO₄, 1.8 mM K₃PO₄ and 2.7 mM KCl, pH 7.4) three times. Then the pellet was air-dried and resuspended in a sample buffer containing 8 M urea, 4% w/v CHAPS and 2% v/v IPG buffer (Amersham Bioscience, UK). The sample was then sonicated, using a Soniprep 150 (Sanyo, UK) fitted with an exponential probe, on ice for 3×10 s. The protein mixture was centrifuged at $14\,000 \times g$ at 4°C for 30 min and the supernatant transferred to fresh microcentrifuge tubes, where the protein extracts were purified using the PlusOne 2-D Clean-Up kit (Amersham Bioscience) according to the manufacturer's directions. Purified protein extract was then quantitated using the PlusOne 2-D Quant kit (Amersham Bioscience) according to the manufacturer's specifications.

2.3 2-DE

Proteins were separated by electrophoresis in the first-dimension with a Multiphor™ II IEF system (Amersham Bioscience) using 18-cm Immobiline dry strips (Amersham Bioscience) with a linear pH range, as described by the manufacturer. Various pH ranges of IPG strips were used and rehydrated with 100 to 120 µg of protein in 350 µL of rehydration solution (8 M urea, 4% w/v CHAPS, 2% v/v IPG buffer, 2.8 mg/mL DTT and 0.002% w/v bromophenol blue)

for at least 16 h at room temperature. IEF was done in three steps consisting of 500 V for 1 min, a linear gradient of 500–3500 V for 90 min, and 3500 V for 6.5 h at 20°C. After focusing was completed, the strips were equilibrated with buffer containing 6 M urea, 130 mM DTT, 30% v/v glycerol, 50 mM Tris-Cl pH 8.8, 2% w/v SDS, and 0.002% w/v bromophenol blue for 15 min, and then with buffer containing 6 M urea, 135 mM iodoacetamide, 30% v/v glycerol, 50 mM Tris-Cl pH 8.8, 4% w/v SDS, and 0.002% w/v bromophenol blue for 15 min. The strips were then placed on wet filter paper for 10 min and applied to SDS 14% w/v self-cast polyacrylamide gels (200 × 160 × 1.5 mm), and electrophoresis was carried out in a Protean II xi cell (Bio-Rad, UK) at 40 mA per gel for 4 h at 10°C with 250 mM glycine, 25 mM Tris-Cl and 0.1% w/v SDS as the running buffer. Low-molecular-weight markers (Sigma, UK) were applied next to the acidic end of the IPG strips.

2.4 Visualization and analysis of the protein spots

Gels were fixed in 50% v/v methanol, 12% v/v acetic acid for 1 h followed by staining with colloidal CBB stain (20% v/v methanol, 10% w/v (NH₄)₂SO₄, 2% v/v phosphoric acid, and 0.1% w/v CBB G-250). After 24 h, the gels were washed several times with water and scanned using a GS-710 calibrated imaging densitometer (Bio-Rad). Analysis of the gels, including protein spot detection and quantitation, was done with PDQuest™ V6.2 software (Bio-Rad). The background was subtracted and the intensities of the image pixels inside the boundary of each protein spot were quantified. Gels were normalized based on the total density in each gel of the matchset. A reference pattern was constructed and each gel in the matchset was matched to the reference gel. Numerous proteins that were uniformly expressed in all patterns were used as landmarks to facilitate rapid gel matching. Differentially expressed proteins were selected according to the Wilcoxon two-sample test ($p < 0.05$), i.e. the protein spot must be more intense (or less intense) on at least six gels in each matchset [13].

2.5 In-gel tryptic digestion and protein identification

The protein spots of interest were excised with a clean scalpel into ~1-mm cubes and each transferred to a 1.2-mL siliconized microcentrifuge tube, washed three times for 30 min in 100 µL of 100 mM NH₄HCO₃ and 60 µL of ACN at room temperature and then incubated in 100 µL ACN for at least 5 min at room temperature. The ACN was removed via aspiration and gels were dried in a centrifugal evaporator for 30 min. The dried gel pieces were hydrated with 25 µL of 50 mM NH₄HCO₃ containing 20 µg/mL of sequencing grade trypsin (Promega, UK) on ice for 30 min and then 30 µL of 50 mM NH₄HCO₃ added (enough to cover the gel pieces) and incubated at 37°C overnight. The peptides extracted after overnight incubation were recovered via aspiration, and additional peptides were recovered from the

gel by extracting with 50% v/v ACN with 5% v/v TFA for 30 min and 83% v/v ACN with 0.1% v/v TFA for a further 30 min. The extracted peptides were recovered via aspiration, dried in a centrifugal evaporator, and suspended in 10 µL of 0.1% v/v TFA. Samples were desalted using C18 ZipTips (Millipore, UK). First, the Zip Tips were rehydrated using 10 µL of 50% v/v ACN, and equilibrated with 10 µL of 0.1% v/v TFA. The samples were each loaded onto a ZipTip by pipetting the peptides extract up and down ten times, and then washed with 10 µL of 0.1% v/v TFA three times. The tryptic fragments were then eluted using 50% v/v methanol, 0.1% v/v acetic acid. Mass analysis was performed on an LCQ Advantage IT instrument equipped with a static nanospray source (Thermo Electron, UK). Peptide samples were each loaded into an EconoTip emitter (Presearch, UK). Static nanospray was performed at the following conditions: positive ionisation mode; spray voltage, 1.8 kV; capillary voltage, 28 V; capillary temperature, 180°C. Data was collected in the full scan and data-dependent MS/MS modes; three microscans were performed, with the maximum ion injection time of 200 ms. In the full-scan mode, ions were collected in the m/z range 300–2000 and the MS/MS collision energy was set to 35%. Data analysis obtained from the MS was performed using the TurboSEQUEST™ program in the BioWorks 3.2 software suite (Thermo Electron). The following parameters were used when creating Dta files from raw files: precursor mass tolerance 1.4 m/z ; minimum number of different ions 15; maximum number of intermediate scans 25; minimum number of grouped scans 1; minimum total ion count 5×10^5 , parent ion mass range 300–2000 Da. A database containing the complete genome sequences of all sequenced streptococcal species (*S. agalactiae* NEM316, *S. pyogenes* MGAS8232, *S. agalactiae* A909, *S. pyogenes* MGAS5005, *S. pyogenes* MGAS6180, *S. thermophilus* CNRZ1066, *S. thermophilus* LMG 18311, *S. pyogenes* MGAS10394, *S. mutans* UA159, *S. agalactiae* 2603V/R, *S. pyogenes* SSI-1, *S. pyogenes* MGAS315, *S. pneumoniae* R6, *S. pneumoniae* TIGR4, *S. pyogenes* M1 GAS) was created by downloading, on 7th April 2006, the appropriate sequences from the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/>). Proteins were then identified by using the Dta files to search the database according to the following parameters: fragment ion mass tolerance of 0 Da; peptide mass tolerance of 1 Da; average mass; maximum number of internal cleavage sites of 2. We evaluated the results from TurboSEQUEST™ using the following filter parameters: proteins must be from the SF370 strain of the M1 serotype of GAS; and peptides must have a cross correlation score (X_{corr}) >1.5 for singly charged peptides, >2.0 for double-charged peptides, and >.5 for triple-charged peptides, with at least two peptides successfully matched; and/or must have a difference in normalised cross correlation score (ΔC_n) >0.2 with at least two peptides successfully matched [14]. The best hit (according to the peptide probability value) as provided by TurboSEQUEST™ and that which fulfilled the filter parameters was then used for protein characterisation.

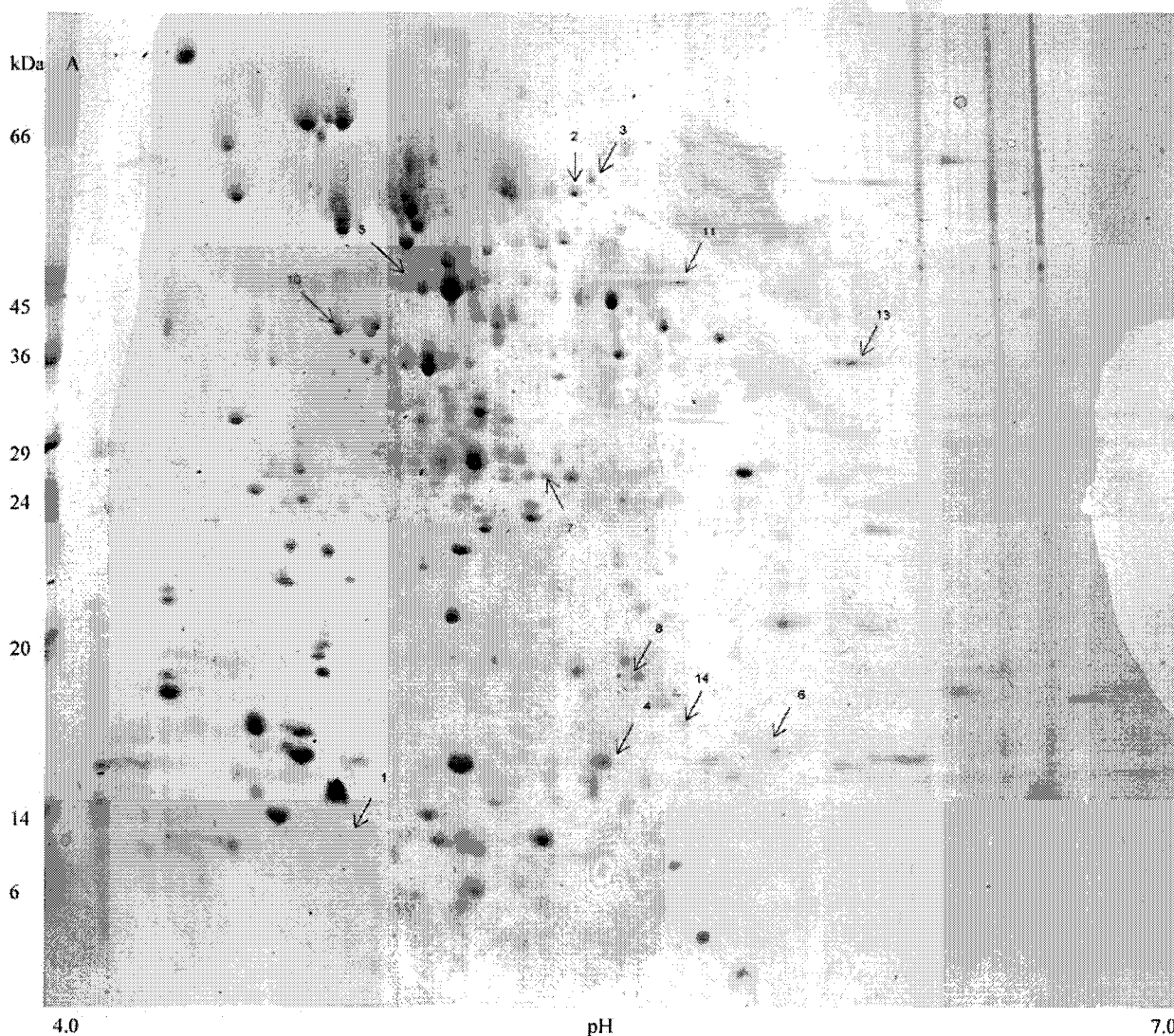
2.6 Further functional analysis of peptides and proteins

TSSview™ v1.0 (Turbinia, UK) was used to identify regions of evolutionary conservation in proteins as determined by high-sensitivity data-mining conducted on 9th May 2006 on ~ three million UniProt entries and their corresponding coding and non-coding regions in each of the six reading frames. Peptides that fulfil the TurboSEQUEST™ X_{corr} and ΔC_n values stated in Section 2.5 were analysed by TSSview™ (including those peptides not from the SF370 strain of the M1 serotype of GAS). Observed/expected (O/E) values $\leq 1.0 \times 10^{-4}$ were considered significant with the level of probability varying according to signature peptide length. The full-length sequences of all identified proteins were also analysed by TSSview™. O/E values $\leq 1.0 \times 10^{-4}$ were considered as highly significant.

3 Results and discussion

3.1 Identification of differentially expressed proteins using 2-DE

Six pairs of 2-DE gels were analysed. The proteins present on each of these gels were isolated from independently grown cultures. An average of 210 ± 20 and 179 ± 20 spots of cell-associated proteins from bacteria cultured in HA enriched media and control media, respectively, were detected by PDQuest™ software on each 2-D gel (Fig. 1). About 80% of the protein spots were matched between the two sets of gels and the correlation coefficient between the two patterns was 0.815. The 18 protein spots highlighted in Fig. 1 represent those that were up- or down-regulated according to the Wilcoxon two-sample test [13]. Among the 18 proteins, 12 proteins were up-regulated or were present only on the 2-D gels



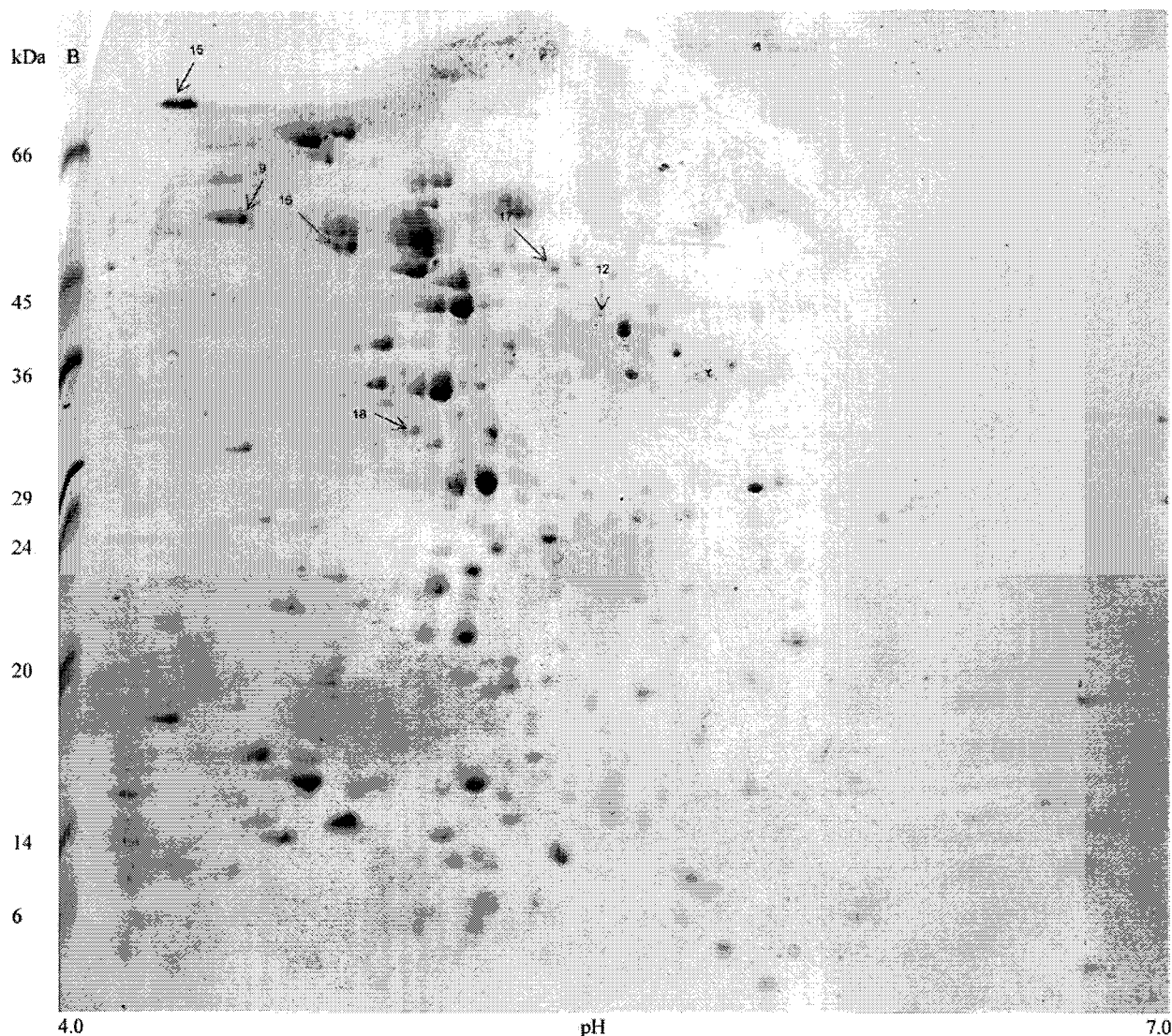


Figure 1. 2-D gel analysis of cell-associated proteins from GAS. (A) Cells cultured in HA-enriched media. The numbers indicate the proteins that were up-regulated in HA-enriched media. (B) Cells cultured in control medium. The numbers indicate the proteins that were down-regulated in HA-enriched media.

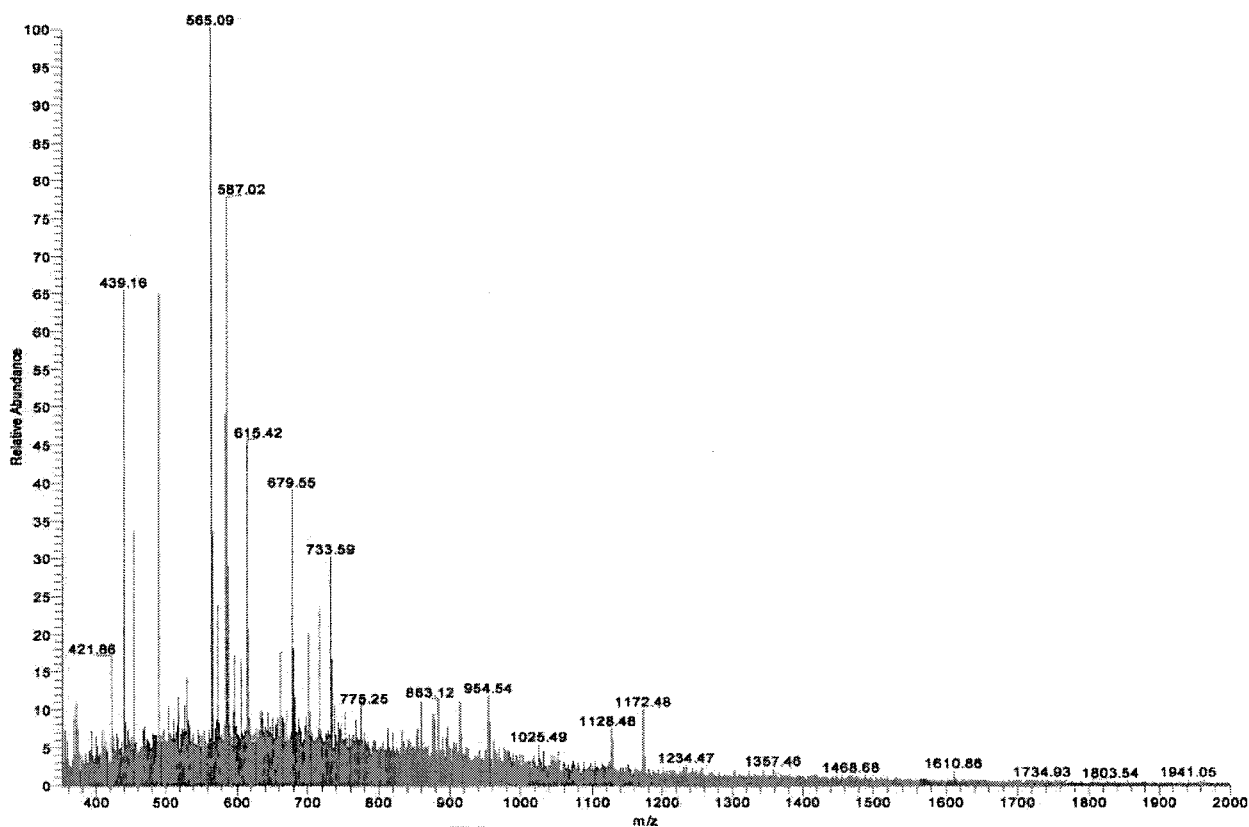
prepared from cells grown in HA-enriched media and 6 proteins were down-regulated or were absent only on the 2-D gels prepared from cells grown in HA-enriched media. The spots corresponding to those 18 proteins were excised from gels and subjected to in-gel trypsin digestion. The resulting peptides were analysed by MS.

3.2 Identity of the differentially expressed proteins using MS

An example of a typical full scan of a protein mass spectrum, the 89.46-kDa putative phosphotransferase system phosphohistidine-containing protein, and the MS/MS spectrum of a

precursor ion (m/z ratio of 489.63) of this protein can be seen in Fig. 2. Of the 18 proteins analysed, spots 1–11 were identified upon searching the streptococcal genome sequence database with the experimental MS/MS spectra using TurboSEQUEST™ (Table 1). Sequence, charge state and filter parameter values for individual peptides from the identified proteins are shown in Table 2. The remaining 7 proteins could not be identified by TurboSEQUEST™ as they did not satisfy the filter parameters defined in the Section 2.5. Three of these protein spots (spots 14, 16 and 17) gave poor quality MS data and therefore could not be analysed further. However, the remaining four proteins spots (spots 12, 13, 15 and 18) were analysed using detailed bioinformatic screening via

spot1 #2-347 RT: 0.06-9.96 AV: 170 NL: 1.12E7 T: + cNSI Full ms [250.00-2000.00]



#4-4 RT:0.09-0.09 NL: 1.38E6

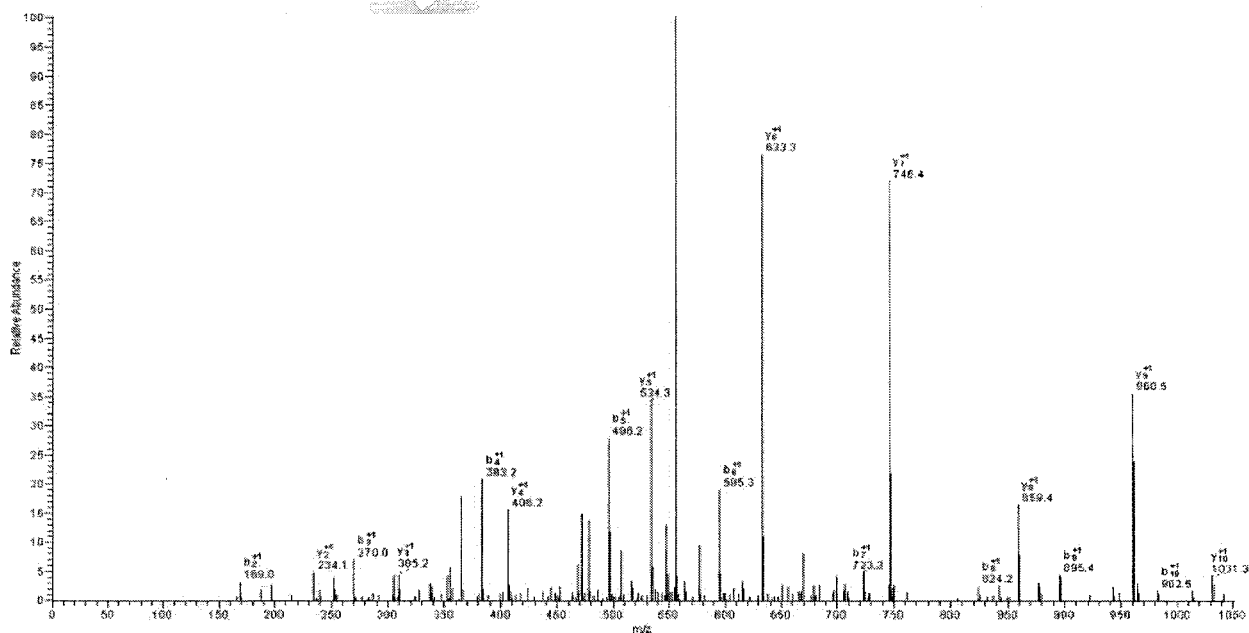


Figure 2. Static nanospray MS. (A) Mass spectrum of a tryptic digestion of protein spot No.1, which was identified as a putative phosphotransferase system phosphohistidine-containing protein using an LCQ Advantage IT MS. (B) MS/MS spectrum of the precursor ion of the protein with an m/z ratio of 489.63.

Table 1. Identity of differentially expressed proteins from the M1 serotype of GAS, strain SF370

Spot no. ^{a)}	Protein ID	Accession ID ^{b)}	M_r/pI ^{c)}	Matched peptides ^{d)}	Protein coverage (%) ^{e)}	Alteration in HA-enriched media	Level of expression (OD area $\times 10^6$) ^{f)}	
							in HA-enriched media ^{g)}	in control media ^{g)}
1	Putative phosphotransferase system phosphohistidine-containing protein ^{h)}	gi 15675305	8946.6/4.74	4	41.38	Up-regulated	124.0 (215.0)	5.0 (15.0)
2	Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase ^{h)}	gi 15675303	50338.4/5.06	2	8.84	Up-regulated	392.1 (347.2)	6.0 (14.9)
3	M1 protein precursor ^{j)}	gi 311758	54203.6/6.53	2	13.22	Up-regulated	127.1 (75.2)	0.0 (17.0)
4	Hypothetical protein SPY1143 ^{i, j)}	gi 15675120	21620.7/5.71	2	14.80	Up-regulated	207.8 (226.4)	23.1 (52.5)
5	Histidine protein kinase ^{h, j)}	gi 15674919	30833.2/5.86	3	15.27	Up-regulated	220.5 (261.5)	16.8 (46.9)
6	tRNA modification GTPase ^{i, j)}	gi 15675063	50509.4/4.67	2	6.77	Up-regulated	91.5 (47.7)	0.0 (22.7)
7	Hypothetical protein SPY1262 ^{h)}	gi 15675451	19944.4/4.93	7	45.25	Up-regulated	130.0 (42.5)	14.5 (21.0)
8	50 S ribosomal protein L10 ^{h)}	gi 15675064	17555.4/5.19	5	30.72	Up-regulated	111.5 (86.3)	35.6 (13.6)
9	Rop A ^{h)}	gi 13622926	47096.3/4.41	5	14.52	Down-regulated	0.0 (47.1)	136.0 (121.2)
10	Collagen-like surface protein ^{j)}	gi 15675773	36260.7/6.91	2	8.62	Up-regulated	409.7 (304.0)	96.1 (10.4)
11	Putative carbamoyl-phosphates synthase small subunit ⁱ⁾	gi 15674869	39757.9/5.57	2	7.78	Up-regulated	132.4 (40.9)	0.0 (0.0)
12	Transcriptional repressor CodY ^{i, k, l)}	gi 15675618	28633.7/4.85	1	7.31	Down-regulated	25.2 (48.0)	105.5 (194.9)
13	Putative formate-tetrahydrofolate ligase ^{i, k)}	gi 15675177	59531.1/5.69	2	5.94	Up-regulated	189.1 (59.1)	0.0 (0.0)

a) Spot number refers to the proteins labelled in Fig. 1.

b) Accession ID of each protein is the GenInfo number in the NCBI protein database.

c) M_r and pI as calculated by PEPTIDEMASS [54].

d) Number of nonredundant peptides identified for each protein.

e) Percent amino acid coverage of entire protein.

f) Sum of the intensities of the image pixels inside the boundary of the spot $\times 10^6$.

g) The value shown is the median of the dataset with the range in parentheses.

h) Proteins identified according to an X_{corr} of >1.5 for single-charged peptides, >2.0 for double-charged peptides, and >2.5 for triple-charged peptides.

i) Proteins identified according to a $\Delta C_n > 0.2$ with at least two peptides successfully matched.

j) Proteins where the theoretical M_r and pI differed significantly from their experimental M_r and pI .

k) Proteins identified by TSSview™.

l) Although functional assignment was based on a single peptide only the likelihood that this peptide was derived by chance alone from another functional group was $O/E = \infty$, i.e. infinitely unlikely.

TSSview™. Protein spot 12, that did not satisfy the filter parameters as only one peptide was successfully matched, was added to the list of identified proteins (Table 1) as the successfully matched peptide was shown using TSSview™ to come from a highly-conserved region of the transcription repressor CodY (O/E ratio: ∞). This single peptide is therefore highly significant for assigning function. Protein spot 13, that did not satisfy the filter parameters as it is from *Streptococcus mutans* and not *Streptococcus pyogenes*, was also added to the list of identified proteins (Table 1) as the two successfully matched peptides were shown using TSSview™ to come from highly-conserved regions of a putative formate-tetrahydrofolate ligase (O/E ratio: ∞). A BlastPAM30 search of these peptides confirmed that highly-identical peptides are present in the same protein from the SF370 strain of the M1 GAS serotype (Table 2). The remaining 2 proteins (spots 15 and 18) that did not satisfy the filter parameters could not be

identified using TSSview™ and BlastPAM30 because the peptides from these proteins are located in conserved regions of proteins that are not present in any homologues of the SF370 strain of the M1 serotype of GAS.

3.3 Transport and biomass/energy production

We observed the up-regulation of proteins in HA-enriched media that are involved in (i) protein and DNA synthesis, i.e. 50 S ribosomal protein L10, the tRNA modification GTPase, a putative formate-tetrahydrofolate ligase and the carbamoyl-phosphate synthase small subunit, (ii) transport, i.e. the phosphohistidine-containing protein, HPr, a component of the phosphotransferase system (PTS) [15] and (iii) glycolysis, i.e. glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Table 1). GAS grown in HA-enriched media reached higher optical densities compared to that grown in control media,

Table 2. Sequence, charge state and filter parameter values for peptides from identified proteins

Spot no.	Protein	Peptides sequence	Charge	X_{corr}	ΔC_n
1	Putative phosphotransferase system phosphohistidine-containing protein	K.DFHIVAETGIHAR.P	3	5.070	0.55
		K.FASDITLDYK.G	2	3.198	0.694
		R.PATLLVQTASK.F	2	3.814	0.602
		K.FASDITLDYKGGK.A	3	2.768	0.549
		K.FASDITLDYKGGK.A	2	3.714	0.707
2	Putative NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	R.PATLLVQTASK.F	1	2.189	0.541
		K.NIVAGAFGYSGQR.C	2	2.847	0.685
		K.AFGIAEQLEVGVTHLNNK.T	2	2.043	0.348
3	M1 protein precursor	K.ALELAIDQASQDYNR.A	2	1.245	0.338
		K.ARLNAMEVAGRDFK.R	2	1.118	0.366
4	Hypothetical protein SPY1143	R.AVILRQGAITK.E	2	1.535	0.255
		K.TGPLIGPSANLSGKASGR.V	2	1.603	0.428
5	Histidine protein kinase	K.DLVFLLRHFEK.A	2	1.000	0.581
		R.HFEKAVRNPLAHLK.P	2	1.000	0.664
		R.REPFYFDDMNAIK.K	2	0.527	0.357
6	tRNA modification GTPase	R.VDLTQAEAVMDIIR.A	2	1.945	0.250
		K.TGTIIDEVMVSVMLAPK.T	3	2.047	0.248
7	Hypothetical protein SPY1262	R.MTDLDVIEVNVK.V	2	4.304	0.600
		K.LVNTESVR.D	2	1.882	0.480
		R.GQLTYDDK.V	2	2.516	0.496
		R.GQLTYDDKVIEK.I	3	1.538	0.204
		K.SIVEEEVKR.M	2	1.292	0.433
		R.STSEFTSHQVENVK.A	2	3.243	0.557
		K.DLTSAIR.G	2	0.648	0.315
		K.KAEQVELIAEK.M	2	4.582	0.565
8	50S ribosomal protein L10	R.GLTVDQDTVLR.S	2	2.122	0.422
		K.AAASIVIVDSR.G	2	3.835	0.693
		R.SLRESGVEFK.V	2	2.052	0.417
9	RopA	K.VINDFTK.T	1	1.181	0.455
		K.AVEVITSTASVK	2	2.313	0.645
		R.VKTNLVIEIAK.A	2	2.409	0.574
		K.GKEWTLAEVVTKEPK.L	3	4.790	0.570
		R.GVITFTISQDK.I	2	1.750	0.523
		K.TNLVIEIAK.A	2	2.674	0.361
		R.GETFAQGPVGPQGEK.G	2	1.490	0.222
10	Collagen-like surface protein	R.GEQGIQKGAGEKGER.G	2	1.604	0.229
		K.GIPGISGIDTRALK.I	2	1.439	0.252
11	Putative carbamoyl-phosphates synthase small subunit	K.MTFGHRGFNHAVR.E	1	1.330	0.275
12	Transcriptional repressor CodY	K.DTYPGGLTTIPIYGGGMR.L ^{a)}	2	1.826	0.361
13	Putative formate-tetrahydrofolate ligase	K.LVLVTAINPTPAGEGK.S ^{b)}	3	1.685	0.541
		K.STITIGLADALNKIGK.K ^{c)}	2	1.484	0.392

a) TSSview™ O/E ratio of ∞.

b) BlastPAM30 E value of 1.0×10^{-5} .c) BlastPAM30 E value of 1.4×10^{-2} .

but there was not a statistically significant difference in the growth rates (data not shown). It was therefore not surprising that we observed the up-regulation of proteins involved in transport and biomass/energy production. These observations are in line with those of Voyich *et al.* [16] who also observed the up-regulation of many of these proteins following analysis of a phagocytic interaction of GAS with human polymorphonuclear leukocytes.

3.4 Virulence factors

3.4.1 LPXTG motif-containing proteins

3.4.1.1 General remarks

Many of the proteins that were shown to be up-regulated in HA-enriched media were LPXTG motif-containing cell-wall

anchored proteins. Such proteins have been shown to play important roles in pathogenesis, such as resistance to phagocytosis, adherence to plasma and extracellular matrix proteins, and degradation of host proteins [17]. As LPXTG motif-containing proteins are covalently attached to the cell wall, we were surprised to see these proteins in the cell-associated protein fraction, as we did not expect them to be released from the cell wall during protein preparation. However, the sonication step used during protein preparation may have released these proteins from the cell wall, as seen elsewhere by Dryla *et al.* [18]. Alternatively, it is possible that the precursor forms of these proteins were derived from the cytoplasm and not the cell wall.

3.4.1.2 M1 protein

We observed an increase in the intensity of a protein spot in HA-enriched media that was identified as the major GAS virulence factor, i.e. the M1 protein [7], an LPXTG motif-containing protein. M1 proteins play several roles in GAS pathogenesis. They protect the bacteria from phagocytic clearance by interfering with complement-mediated uptake [19], they can attach to keratinocytes and cause skin infection [20], and, during the throat infection, they promote bacteria-bacteria interaction following attachment to tonsillar epithelial cells [21]. Functional analysis by TSSview™ of the sequence of the M1 protein showed that this protein has several conserved regions (Fig. 3). As shown in Fig. 3, several conserved areas of M-like proteins were discerned. M-like

proteins share a similar structure with M proteins and are encoded by *emmL* genes. Both M protein and M-like proteins have been shown to have the ability to bind to the Fc region of immunoglobulin [22–24], and this ability can be further subdivided into those proteins that bind immunoglobulin A or proteins that bind to immunoglobulin G [25–28]. The ability of M1 proteins to bind immunoglobulin plays an important role in anti-phagocytosis and interestingly the conserved regions that facilitate immunoglobulin binding were shown by TSSview™ to be present in the GAS M1 protein sequence (Fig. 3). In addition, the M1 protein sequence was shown to have several conserved regions for plasminogen-binding function (Fig. 3). This function has been shown to facilitate the accumulation of plasmin on the GAS surface, and can therefore help break host barriers and promote migration of GAS, which is important for the invasive GAS infections [29]. The M1 protein also possessed conserved regions for fibrinogen-binding activity (Fig. 3). This activity can precipitate fibrinogen that prevents the killing of GAS in host blood [30]. As all of these M1 protein activities have been experimentally determined, it confirms that TSSview™ is a powerful tool in aiding the functional characterisation of proteins.

3.4.1.3 Collagen-like surface protein

We also observed the up-regulation in HA enriched media of another LPXTG motif-containing cell-wall anchored proteins, i.e. a collagen-like surface protein. Lukomski and



Figure 3. Identity of regions of unusual conservation in the M1 protein from the SF370 strain of the M1 serotype of GAS, as determined by TSSview™. Plasminogen-binding activity, —————; M-like proteins, - - - -; Fibrinogen-binding activity,; Immunoglobulin G binding activity, ————; Immunoglobulin A-binding activity, - · - · - ·; Fc-gamma receptor, — · — · — ·.

Nakashima [31] demonstrated that a collagen-like surface protein mutant possessed a significantly reduced ability to adhere to human epithelial cells grown in culture and was significantly less pathogenic in a mouse model of soft tissue infection, which suggested that collagen-like surface protein is a GAS virulence factor that participates in the adherence of GAS to host cells and soft tissue pathology.

3.4.2 Other virulence factors

Interestingly, the glycolytic enzyme, GAPDH [32, 33] and the transport protein, HPr [34], that we have shown to be up-regulated in HA-enriched media (see Section 3.3), have been identified at the cell surface of pathogenic streptococci, with GAPDH shown to be implicated in pathogenesis [35–37]. It has been demonstrated that GAPDH was able to bind to fibronectin, lysozyme and cytoskeletal proteins and this multiple binding capacity of the GAPDH has been postulated to play a role in the colonization, internalization and the subsequent proliferation of GAS [35]. Additionally, GAPDH produced by GAS has been found to have adenosine diphosphate-ribosylating activity [35]. Since adenosine diphosphate-ribosylation is an important component of intracellular signalling events, GAPDH has been shown to have a role in mediating cell-to-cell communication between GAS and pharyngeal cells [36]. Recently, it has also been found that GAPDH helped GAS to bind to the complement component, C5a [38]. Interestingly, two other proteins, that were up-regulated in HA-enriched media, have been shown to be involved in signal transduction, the transport protein, HPr [37], and histidine protein kinase, which works as a transmembrane sensor detecting extracellular environmental changes to regulate the expression of several GAS virulence factors to enhance survival [39, 40]. These enzymes may therefore be important in facilitating the sensing of the external environment by GAS during infection. However, it should be noted that the sum of surface and cytosolic forms of the enzymes were analysed in this study.

3.5 Down-regulation of two regulators of virulence factors

A regulator of streptococcal virulence factors, RopA, was down-regulated in HA-enriched media (Table 1). RopA is a regulator of streptococcal pyrogenic exotoxin B (SpeB) and assists SpeB in translocation via the secretory pathway and functions as a molecular chaperone to achieve an active conformation of SpeB [41]. The down-regulation of RopA suggests that the level of expression of SpeB may have decreased in our system, yet we have not identified SpeB in our studies, as it is secreted and we have only analysed cell-associated proteins. However, the function of SpeB is uncertain, as it can hydrolyze M proteins and C5a peptidase [42], which are important in anti-phagocytosis. In fact, several studies have shown that increased expression of SpeB resulted in decreased virulence of GAS [43, 44]. Both Kansal *et al.* [45]

and Raeder *et al.* [42] reported the inverse relationship between SpeB production and expression of M protein, and Kansal *et al.* [45] hypothesized that the reduced expression or activity of SpeB may be advantageous to the bacteria in cases of streptococcal toxic shock syndrome.

We also observed the down regulation of a transcriptional repressor, CodY, as was also seen by Graham *et al.* [46] when wild-type GAS was exposed to human blood. Interestingly, Malke *et al.* [47] have shown that CodY is a growth phase-dependent positive transcriptional regulator of Mga, the positive multiple regulator, and CovR, which is a member of the two-component global regulatory system CovRS. The positive regulation of Mga and CovR by CodY positively affect the expression of six virulence factors (hyaluronate synthase A, immunoglobulin G degrading endopeptidase, pyrogenic exotoxin H, C5a peptidase, a cell surface proteinase, and a collagen-like surface protein). However, the same report showed that the regulation of Mga and CovR by CodY has been shown to occur at the mid-exponential phase of cell growth [47], while our cells were harvested at early stationary phase. Interestingly, at early stationary phase CodY has also been shown to act as a guanosine 5-triphosphate-binding protein [48] that senses the intracellular guanosine 5-triphosphate concentration as an indicator of nutritional limitations. It is therefore possible that the observed down-regulation of CodY is due to a reduction in the levels of guanosine 5-triphosphate in our system at early stationary phase.

It should be noted that in addition to the two streptococcal regulators we observed, GAS has a series of global regulators of virulence factors [8]. The previously mentioned CovR, for example is involved in the regulation of the synthesis of the streptococcal HA capsule [49]. Although we did not detect altered expression of CovR in HA-enriched media, further post-genomic analyses will add to our understanding of the contribution of carbohydrate metabolism to GAS pathogenesis.

3.6 Hypothetical proteins

TurboSEQUEST™ identified two hypothetical proteins, the functions of which have not been annotated in the genome database. By analysis of these proteins using TSSview™, we found that hypothetical protein SPY1143 has conserved regions that are also present in SUA5 protein (Table 3). SUA5 is required for the normal growth of yeast cells [50], and similar sequence have been found in the M28 GAS serotype [51]. However, the function of SUA5 is not yet known, therefore it is possible that this protein, due to its up-regulation in our system, is involved in GAS pathogenesis. Another hypothetical protein, SPY1262 has been shown by TSSview™ to include some of the conserved regions of the general stress protein, Gls24 (Table 3). This protein, in *Enterococcus faecalis*, has been found to be implicated in virulence as well as stress response [52, 53] and may therefore have a similar role in GAS. The observed up-regulation of this protein is consistent

Table 3. Functional analysis of two hypothetical proteins using TSSview™

Hypothetical protein	Functionality	Peptide sequence	Conserved regions
Hypothetical protein SPY1143	SUA5	VLPTETVYGLFAKALDEKAV	15–34
	Necessary for normal growth	AFLPGPLTIILKANDQVP	76–93
		TGPLIGPSANLSG	120–132
Hypothetical protein SPY1262	Gls24	DLTSAIRGQLTYDDKVIK	13–30
	General stress protein	GVNVEVGKKQVAVDL	66–80

with the result of the transcriptome study performed by Graham *et al.* [46] who found that SPY1262 was up-regulated in the late stage of GAS exposure to human blood.

4 Concluding remarks

We have used 2-DE in conjunction with static nanospray MS to identify proteins from a GAS M1 serotype that have altered expression as a result of propagation in HA-enriched media. Some of these proteins are major virulence factors of GAS, such as the M1 protein and a collagen-like surface protein, some are regulators of virulence factors, such as RopA, and some are proteins that are not commonly associated with virulence, *e.g.* the glycolytic enzyme GAPDH, but have been shown to play important roles in streptococcal pathogenesis. Interestingly, we also identified a hypothetical protein, which may be involved in GAS pathogenesis. Unlike many models of streptococcal pathogenicity, we have shown that GAS grown in HA-enriched media is readily amenable to proteomic analysis due to the relatively large amounts of biomass propagated. The fact that we have observed the up-regulation of several recognised virulence factors indicates that this simple biological system reflects several elements of GAS pathogenesis in humans and is therefore a valuable tool in the identification of potential virulence factors.

We would like to thank Iain Sutcliffe (Northumbria University) for his helpful comments on the manuscript and John Perry (Freeman Hospital, Newcastle upon Tyne) for his help with storage of *S. pyogenes* SF370. Meng Zhang is a recipient of the Universities UK Overseas Research Students Awards Scheme scholarship.

5 References

- [1] Mims, C., Dockrell, H., Goering, R., Roitt, I. *et al.*, *Medical Microbiology*, Mosby, USA 1998.
- [2] Carapetis, J. R., Steer, A. C., Mulholland, E. K., Weber, M., *Lancet Infect. Dis.* 2005, 5, 685–694.
- [3] Hynes, W. L., Ferretti, J. J., *Infect. Immun.* 1989, 57, 533–539.
- [4] Laurent, T. C., Fraser, J. R., *FASEB J.* 1992, 6, 2397–2404.
- [5] Smith, N. L., Taylor, E. J., Lindsay, A. M., Charnock, S. J. *et al.*, *Proc. Natl. Acad. Sci. USA* 2005, 102, 17652–17657.
- [6] Hynes, W. L., Walton, S. L., *Microbiol. Lett.* 2000, 183, 201–207.
- [7] Cunningham, M. W., *Clin. Microbiol. Rev.* 2000, 13, 470–511.
- [8] Graham, M. R., Smoot, L. M., Lei, B., Musser, J. M., *Curr. Opin. Microbiol.* 2001, 4, 65–70.
- [9] Lei, B., Mackie, S., Lukomski, S., Musser, J. M. *Infect. Immun.* 2000, 68, 6807–6818.
- [10] Thongboonkerd, V., Luengpailin, J., Cao, J., Pierce, W. M. *et al.*, *J. Biol. Chem.* 2002, 277, 16599–16605.
- [11] Chaussee, M. A., Callegari, E. A., Chaussee, M. S., *J. Bacteriol.* 2004, 186, 7091–7099.
- [12] Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J. *et al.*, *Proc. Natl. Acad. Sci. USA* 2001, 98, 4658–4663.
- [13] Houtman, R., Krijgsveld, J., Kool, M., Romijn, E. P. *et al.*, *Proteomics* 2003, 3, 2008–2018.
- [14] Eng, J. K., McCormack, A. L., Yates, J. R. *J. Am. Soc. Mass Spectrom.* 1994, 5, 976–989.
- [15] LiCalsi, C., Crocenzi, T. S., Freire, E., Roseman, S., *J. Biol. Chem.* 1991, 266, 19519–19527.
- [16] Voyich, J. M., Sturdevant, D. E., Braughton, K. R., Kobayashi, S. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, 100, 1996–2001.
- [17] Hynes, W., *Front. Biosci.* 2004, 9, 3399–3433.
- [18] Dryla, A., Prustomersky, S., Gelbmann, D., Hanner, M. *et al.*, *Clin. Diagn. Lab. Immunol.* 2005, 12, 387–398.
- [19] Husmann, L. K., Yung, D. L., Hollingshead, S. K., Scott, J. R., *Infect. Immun.* 1997, 65, 1422–1430.
- [20] Okada, N., Pentland, A. P., Falk, P., Caparon, M. G., *J. Clin. Invest.* 1994, 94, 965–977.
- [21] Caparon, M. G., Stephens, D. S., Olsen, A., Scott, J. R., *Infect. Immun.* 1991, 59, 1811–1817.
- [22] Guyre, P. M., Campbell, A. S., Kniffin, W. D., Fanger, M. W. *J. Clin. Invest.* 1990, 86, 1892–1896.
- [23] La Penta, D., Zhang, X. P., Cleary, P. P., *Mol. Microbiol.* 1994, 12, 873–879.
- [24] Sjobring, U., Ringdahl, U., Ruggeri, Z. M., *Blood* 2002, 100, 4470–4477.
- [25] Frithz, E., Heden, L. O., Lindahl, G., *Mol. Microbiol.* 1989, 3, 1111–1119.
- [26] Gomi, H., Hozumi, T., Hattori, S., Tagawa, C. *et al.*, *J. Immunol.* 1990, 144, 4046–4052.
- [27] Burova, L. A., Nagornev, V. A., Pigarevsky, P. V., Gladilina, M. M. *et al.*, *Appl. Microbiol.* 2005, 113, 21–30.
- [28] Podbielski, A., Schnitzler, N., Beyhs, P., Boyle, M. D., *Mol. Microbiol.* 1996, 19, 429–441.

- [29] Berge, A., Sjobring, U.J. *Biol. Chem.* 1993, **268**, 25417–25424.
- [30] Johansson, H. M., Morgelin, M., Frick, I. M. *Microbiology* 2004, **150**, 4211–4221.
- [31] Lukomski, S., Nakashima, K., Abdi, I., Cipriano, V. J. *et al.*, *Infect. Immun.* 2000, **68**, 6542–6553.
- [32] Pancholi, V., Fischetti, V. A., *J. Exp. Med.* 1992, **176**, 415–426.
- [33] Wilkins, J. C., Homer, K. A., Beighton, D., *Appl. Environ. Microbiol.* 2002, **68**, 2382–2390.
- [34] Dixon, S., Haswell, M., Harrington, D., Sutcliffe, I. C., *Syst. Appl. Microbiol.* 2001, **24**, 486–489.
- [35] Pancholi, V., Fischetti, V. A., *Proc. Natl. Acad. Sci. USA* 1993, **90**, 8154–8158.
- [36] Pancholi, V., Fischetti, V. A., *Adv. Exp. Med. Biol.* 1997, **418**, 499–504.
- [37] Wilkins, J. C., Beighton, D., Homer, K. A., *Appl. Environ. Microbiol.* 2003, **69**, 5290–5296.
- [38] Terao, Y., Yamaguchi, M., Hamada, S., Kawabata, S., *J. Biol. Chem.* 2006, **281**, 14215–14223.
- [39] Senadheera, M. D., Guggenheim, B., Spatafora, G. A., Huang, Y. C. *et al.*, *J. Bacteriol.* 2005, **187**, 4064–4076.
- [40] Musser, J. M., DeLeo, F. R., *Am. J. Pathol.* 2005, **167**, 1461–1472.
- [41] Collin, M., Olsen, A., *Infect. Immun.* 2003, **71**, 2983–2992.
- [42] Raeder, R., Woischnik, M., Podbielski, A., Boyle, M. D., *Res. Microbiol.* 1998, **149**, 539–548.
- [43] Raeder, R., Harokopakis, E., Hollingshead, S., Boyle, M. D., *Infect. Immun.* 2000, **68**, 744–751.
- [44] Ashbaugh, C. D., Warren, H. B., Carey, V. J., Wessels, M. R., *J. Clin. Invest.* 1998, **102**, 550–560.
- [45] Kansal, R. G., McGeer, A., Low, D. E., Norrby-Teglund, A., Kotb, M., *Infect. Immun.* 2000, **68**, 6362–6369.
- [46] Graham, M. R., Virtaneva, K., Porcella, S. F., Barry, W. T. *et al.*, *Am. J. Pathol.* 2005, **166**, 455–465.
- [47] Malke, H., Steiner, K., McShan, W. M., Ferretti, J. J., *Int. J. Med. Microbiol.* 2006.
- [48] Ratnayake-Lecamwasam, M., Serror, P., Wong, K. W., Sonenshein, A. L., *Genes Dev* 2001, **15**, 1093–1103.
- [49] Levin, J. C., Wessels, M. R., *Mol. Microbiol.* 1998, **30**, 209–219.
- [50] Na, J. G., Pinto, I., Hampsey, M., *Genetics* 1992, **131**, 791–801.
- [51] Green, N. M., Zhang, S., Porcella, S. F., Nagiec, M. J. *et al.*, *J. Infect. Dis.* 2005, **192**, 760–770.
- [52] Teng, F., Nannini, E. C., Murray, B. E., *J. Infect. Dis.* 2005, **191**, 472–480.
- [53] Nannini, E. C., Teng, F., Singh, K. V., Murray, B. E., *Infect. Immun.* 2005, **73**, 7772–7774.
- [54] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. *et al.*, in: Walker, J. M. (Ed.), *The Proteomics Protocols Handbook*, Humana Press, Totowa, NJ 2005, pp. 571–607.